DOCUMENT RESUME

ED 051 547 EA 003 500

AUTHOR	Jung, Steven M.; And Others
TITLE	Study of the Use of Incentives in Education and the Feasibility of Field Experiments in School Systems. Final Report.
INSTITUTION	American Institutes for Research in the Behavioral Sciences, Palo Alto, Calif.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	AIR-2027-1-71-FR
PUB DATE	29 Jan 71
CONTRACT	OEC-0-70-5035
NOTE	152p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	Bibliographies, Data Analysis, *Educational Accountability, Educational Experiments, *Feasibility Studies, Feedback, *Incentive Systems, Input Output Analysis, Learning Theories, Literature Reviews, *Motivation, Multiple Regression Analysis, Performance Contracts, *Reinforcement, Social Reinforcement

ABSTRACT

This document reports the results of a 6-month investigation into the possibility of instituting operational incentive systems in schools. An experiment is suggested that would examine the effects of a wide range of monetary and nonmonetary rewards on students, teachers, administrators, and parents based on demonstrated gains in student performance objectives in reading and mathematics. Considerable attention is focused on the collection of process and output information and the suggested use of existing school records, self-report instruments, indepth interviews, behavioral observations, and functional-level paper and pencil tests. The basic structure of a multiyear experimental program encompassing several sites is presented, with suggestions for potential local, State, and Federal funding. Appendixes include documentation of the EZ Sort files, documentation of several current projects that make use of incentives to students, and letters of interest from school districts that were contacted as part of the feasibility study. (Author)

AIR-2027-1/71-FR

ED051547

FINAL REPORT

Contract No. OEC-0-70-5035

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

STUDY OF THE USE OF INCENTIVES IN EDUCATION AND THE FEASIBILITY OF FIELD EXPERIMENTS IN SCHOOL SYSTEMS

Steven M. Jung
Dewey Lipe
Peggy S. Wolfe

American Institutes for Research
Box 1113
Palo Alto, California 94302

January 29, 1971

Submitted to:
U.S. Department of Health, Education, and Welfare
U.S. Office of Education
Office of Program Planning and Evaluation

EA 003 500

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official U.S. Office of Education position or policy.

ABSTRACT

Reported herein are the results of a six month investigation into the possibility of instituting operational incentive systems in schools. On the basis of an extensive review of current literature, which is included, an experiment is suggested to examine the effects of a wide range of monetary and non-monetary rewards to students, teachers, administrators, and parents. These rewards would be based upon demonstrated gains on student performance objectives in reading and mathematics. Problems in experimental design and data analysis, inherent in such field experiments, are discussed; partial resolutions, utilizing descriptive data in a quasi-experimental time series paradigm, replication of effects across sites, and multiple regression analysis, are suggested. Considerable attention is accorded to the details of collecting process and output information, suggesting the use of existing school records, self-report instruments, in-depth interviews, behavioral observations, and functional-level paper and pencil tests. The application of such measurement procedures would in itself add incentives to the educational process. The basic structure of a multi-year experimental program encompassing several sites is presented, with suggestions for potential local, state, and federal funding.

Appendices to the report include documentation of the EZ Sort files which were established as part of the literature review, documentation of several current projects which make major use of incentives to students, and letters of interest from school districts which were contacted as part of the feasibility study.

TABLE OF CONTENTS

	page
Abstract	i
Introduction	1
Review of Literature on Educational Uses of Incentives	4
Theoretical Rationales	4
Types of Incentives	6
Target Populations	20
Modes of Incentive Delivery	24
Timing of Reinforcements	34
Criterion Measures for Incentive Delivery	37
Monitoring Instructional Processes	45
General Design Considerations	52
Student Incentive Models	56
Students as Recipients--Microincentives Model	56
Students as Recipients--Macroincentives Model	58
School Personnel Incentive Models	61
Teacher Incentives Models	62
Administrators as Recipients	65
Parent Incentive Model	67
Summary of Models	68
An Incentives Experiment	70
Possible Organization of a School-Based Experiment	70
Experimental Options	73
Project Cost Estimates and Continuation Options	77
Bibliography	80
Appendices	91
A. Documentation of Educational Programs Utilizing Incentives to Students	93
B. Documentation of EZ Sort Literature Retrieval System	123
C. PERT Networks for Three Basic Incentives Models	127
D. Letters of Interest from Contacted School Districts	141
E. Short Introduction to Incentives Experiments	147

INTRODUCTION

The use of incentives in educational practice is universal. Incentives, whether or not they are identified as such, exist for all participants in the educational process. In the present context, incentives are thought of as identifiable consequences of behavior which act to guide the future form and frequency of that behavior. Such factors as money, security, knowledge of personal success, peer or authority figure approval, fear of failure, etc., are probably operating to influence a large percentage of the behaviors which could be observed and classified in any school in the country. In this sense, the educational enterprise is not unlike other forms of human enterprise.

Recent events have stimulated serious interest in the use of incentives to improve academic performance. One such event is the flurry of contracts between school systems and private firms which bind the latter to produce specified reading and mathematics achievement gains in students in order to be paid for instructional services rendered (Education Turnkey Systems, 1970), giving rise to the notion that outside firms know something that school personnel do not know about causing students to learn. Central to these events is the belief that the educational programs of the past decade have not produced impressive results and have especially failed the so-called "deprived" student. Whereas these failures have produced a certain pessimism in some circles, other educators have thought enough of the power of currently available techniques to venture their own capital on a guaranteed-performance-or-no-pay basis. An examination of these techniques ususally reveals a heavy emphasis on technological innovations and "incentives" to learners.

The first portion of this report examines research evidence regarding the effectiveness of various types of incentives in improving student performance, using various modes and schedules of incentive delivery, directed toward various identifiable incentive recipients. This evidence provides a basis for the formulation of a school-based experiment which could be done to investigate conditions where positive incentives are <u>identified</u> and <u>manipulated</u> by an outside party <u>versus</u> the effect of conditions where incentives exist "in their natural state." The ultimate goal of this experiment

would be to determine the comparative benefits of positive incentives apart from the influence of major revisions in the school curriculum. In essence, the proposed study would investigate the effects of predetermined stipulations on the outcomes of the educational process, as opposed to changes in the materials or procedures which constitute the process. It is postulated that modifying consequences based on outcomes would lead to or require changes in methods. These changes would be of major interest; indeed, one of the major purposes of the experiment would be to gather reliable "process" information regarding their nature.

The experiment would be planned to extend over the course of an entire school year. Previous experience suggests that experiments of shorter duration cannot be counted on to produce replicable results. Moreover, serious consideration should be given to the notion of studying several sites on a longitudinal basis for several years. It is felt that the importance of the potential findings of such experiments well justifies the attempt to study their effects in depth. It is quite likely that incentive effects are cumulative, and the differences between organized incentive and unspecified incentive conditions would become more apparent over the course of several years. Some researchers (e.g., Clark, 1970; Green & Stachnik, 1968; Staats, Minke, Goodwin, & Landeen, 1967; Tyler, 1969) have noted that initial failure experiences produce a continuous cycle of deepening student failure. It is possible that providing identifiable incentives could cause students to gain the skills that would allow them to experience more success in school and hence become more receptive to the influence of less tangible incentives which already exist in the classroom situation. These effects should become observable in longitudinal data collected after external incentives have been phased out.

The most crucial requirement for implementing the experiment proposed here is that outputs of the educational enterprise must be made a matter of record, subject to external scrutiny. The often used term "accountability" (Lessinger, 1970) contains the appropriate flavor. As long as teaching and learning take place in a vacuum, incentives will be extremely difficult to direct in other than a haphazard fashion. On the other hand, when the expected outputs are identified in advance and well known as such by all participating parties, incentives will follow almost as a matter of course (cf. Wynne, 1971).

The proposed experiment is organized according to the diverse target populations which could receive incentives. Seven separate experimental models are described utilizing these target populations. The proposed experiment would contain a range of these models. Only populations which can be considered as legitimate entities having major responsiblity for the educational process are identified. These include students, teachers, administrators, and parents. Outside profit-based organizations are omitted.

In proposing a multi-site field experiment on the effects of delivering incentives contingent upon gains in student achievement, the authors emphasize that such a study would push experimental methodology to the threshold of its developed capability, perhaps beyond. However, there are important questions which cannot be answered without such an attempt. Basically, the questions are these: can educational goals for students be identified which are significant educationally yet provide suitable criteria for the assignment of incentives; if so, will major responsible parties in the educational process <u>allow</u> external incentives to be made contingent upon the attainment of such goals; if so, can incentives be <u>effective</u> in improving goal attainment over and above the success of existing educational programs; if so, how was this accomplished, and if not, why not?

It is interesting to note that the first two questions above must each be answerable in the affirmative before experimental methods can possibly be brought to bear to answer the latter questions. And only by <u>embarking</u> on an experiment can these first questions be posed. The project staff thus have outlined a set of procedures which could be used to establish measurable criteria of student achievement. Next a range of possible models was constructed which could provide and evaluate the effects of incentives. On the basis of these formulations, a selected sample of school districts around the country was contacted to determine potential interest. In view of the magnitude of the problem involved, the preliminary responses from these districts were quite encouraging. Several indicated an intent to proceed on their own in establishing operational incentives systems. Goals and models were modified according to some of the concerns of these involved educators and project consultants. Finally, a complete experimental program along with potential funding sources was explored. The authors are hopeful that this report can serve as a vehicle to bring about an operational version of such a program.

REVIEW OF THE LITERATURE ON THE EDUCATIONAL USES OF INCENTIVES

The purpose of this review is to summarize the research on the use of incentives to influence student performance, identifying demonstrated principles. For an orderly presentation that covers the range of possibilities for incentive manipulation, the topic will be organized to focus on theoretical rationales, types of incentives, target populations, modes of incentive delivery, and timing of reinforcements.

Theoretical Rationales

Two classes of theories are relevant to motivation of human performance. One class of theories regards the internal state of the organism. These theories hypothesize constructs about what happens within a human organism to direct and energize its behavior. Theories of personality, emotion, and homeostatic states have been employed to describe the internal mechanism of striving to learn. McClelland (1951), for example, hypothesized a "need for achievement" that varies among individuals and accounts for some of their performance differences in a given educational environment. Such traits are difficult to measure reliably, and, being relatively enduring characteristics of the individual, are of educational importance only if instructional treatments can be found which interact positively with them. Evidence for such attribute-by-treatment interaction is quite elusive at present (cf. Cronbach & Snow, 1969).

The other class of theories focuses on events outside the organism; those events may be described as incentives to influence the organism's performance. According to Bandura (1969):

> Incentive theories of motivation assume that behavior is largely activated by anticipation of reinforcing consequences. From this point of view, motivation can be regulated through arrangement of incentive conditions and by means of satiation, deprivation, and conditioning operations that affect the

> relative efficacy of various reinforcers at any given time. Thus, for example, in producing intellectual strivings in children who display little interest in academic pursuits, one would arrange favorable conditions of reinforcement with respect to achievement behavior rather than attempt to create in some ill-defined way an achievement motive, the presence of which is typically inferred from the behaviors it presumably actuates. [pp. 226-227]

Much of the current research designed to organize external events as incentives for academic skill acquisition owes its impetus to the research of Arthur W. Staats (e.g., Staats, 1957; Staats, Staats, Schutz, & Wolf, 1962; Staats, Finley, Minke, & Wolf, 1964). More recent evidence that so called "motivation to learn" can be developed in children by restructuring the educational environment comes from demonstrated success with two strikingly unmotivated groups: convicted and sentenced delinquents and non-verbal autistic children. Autistic children have been trained to talk (Lovaas, 1968) and delinquent adolescents with a long history of school failure have improved their reading and achievement test scores (Cohen, 1968) through the use of environmental manipulation techniques.

This point of view about learning, first exemplified on a large scale by the programmed instruction movement, is that failure of a learner to achieve the instructional objectives of a program reflects flaws in the program, not an inadequacy of the learner. Such perspectives on human learning are beginning to be generalized to our educational system (cf. Coleman, 1969) and even to our very cultural structure (cf. Skinner, 1966a; Schwitzgebel, 1970). The notion here is that motivation to learn, or the lack of it, is a behavioral response to environmental contingencies rather than a manifestation of sloth, laziness, or recalcitrance. If a student in a certain educational setting does not achieve the desired objectives, this means that the setting is so structured that the student is unable to learn. It also means that the setting could be restructured so that the same student would achieve the educational objectives deemed important for him.

Those aspects of learning theories that concern the relationships of stimuli and behavioral consequences to future performance are highly relevant to incentive manipulation of performance. No one theory, however, will provide a complete design for planning the most effective use of incentives to induce learning achievement. No one theory, in fact, has an absolute claim on incentive manipulation. Operant learning theories are perhaps the most

relevant. They include the giving or withholding of immediate rewards and punishments. Contingent placement of incentives, essential in operant conditioning, however, is not always required to motivate learning. As will be shown, humans have amply demonstrated the ability to provide cognitive mediation between behaviors emitted now and rewards promised much later.

Types of Incentives

Many of the incentive theories of motivation have been derived from experimental psychology, and it is not surprising, therefore, that some of the same incentives used in laboratories were also tried in classrooms. Most such incentives may be regarded as stimulus manipulation. Discussed under stimulus manipulation will be material incentives, social incentives, knowledge of results, secondary reinforcers such as tokens, points, or money, and aversive stimuli. Another class of incentives, however, has evolved from the theoretical formulation of Premack (1965). These and other related incentives will be grouped under response manipulation. The topic of response manipulation will include the use of high probability responses to reward desired low probability responses, and it will also include a discussion of students' self-management of learning incentives.

Stimulus Manipulation

Material Incentives. The use of material incentives for learning has a long and vigorous history in experimental psychology. Food, in particular, has been popular because deprivation level can be easily controlled. Fruit and cookies or sandwiches were given as reinforcers to Black children in an experimental preschool located in a depressed area of Kansas City (Risley & Hart, 1968). The experiment sought to develop correspondence between what the children said they had done and what they actually did. Snacks given at the regular snack time but made contingent on desired behavior were successful incentives.

Chadwick and Day (1970) and Day and Chadwick (1970) used food and other material reinforcers in addition to other types of incentives in a class of

30 Black and Mexican-American students who were all severe behavior problems. The school-furnished lunch was used as a reinforcer along with school store items such as candy, gum, goldfish, clothes, jewelry, etc. In conjunction with the other type of incentives, the use of these material incentives was instrumental in improving both social behavior and academic performance.

Other material incentives have been used in educational settings. Benowitz and Busse (1970) found that a box of crayons made contingent on "doing well" on a spelling test did result in improved spelling scores of fourth grade Black children in two large urban ghetto schools. The teachers simply told their classes each day during the experiment, "If you do very well on your spelling test this week, you will be given a nice prize, a box of crayons."

In some studies, an array of material incentives has been made available contingent on improved performance. In a laboratory learning experiment with 160 Caucasion kindergarten children, materials such as candies, trinkets, small cars, and dolls were made available for successful performance (Marshall, 1969). In that particular experiment, material rewards hindered performance of both groups of children. The author speculated that when children took time to select their reward, they interrupted the continuity of the task, lowering their scores. This same finding was reported by Spence (1966). She found that candy rewards yielded poorer results than verbal statements of right or wrong on a discrimation learning task with preschoolers, second and third graders, and fifth and sixth graders. In another experiment which compared candy vs. teacher praise statements such as "that's good," the two conditions did not differ in their effects on performance of a simple descrimination learning task (Unikel, Strain, & Adams, 1969). Five and six year old White children enrolled in Project Head Start in a rural area were subjects in that laboratory experiment.

Social Incentives. Teacher praise has been the most widely studied of social incentives. Kennedy and Willcutt (1964) reviewed 33 studies performed over the previous 50 years on the use of praise and blame as incentives. Praise has been compared with material incentives, reproof, and knowledge of results. Characteristics of praise givers and praise receivers have been

evaluated for their relationships to the performance of the praise receivers; and, more recently, group social influences have been assessed.

O'Leary and Becker (1969) employed an experimental analysis of behavior procedure to evaluate the effects of praise, ignoring, and reprimands on disruptive behavior of a class of 19 first graders during its rest period. They found that praising appropriate behavior and ignoring disruptive behavior reduced the average time in disruptive behavior from 54% to 32%. When reprimands were reinstated, disruptive behavior increased to baseline level but dropped again to an average of 35% when all praise and no reprimands again were given.

Two other studies of praise in conjunction with some form of disapproval contingent on conforming and not conforming to the established protocol of the classroom further substantiated the O'Leary and Becker results (viz., McAllister, Stackowaik, Baer, & Conderman, 1969; Wasik, Senn, Welch, & Cooper, 1969). The experiment by McAllister et al. was with secondary students and the study by Wasik et al. was with two second grade Black girls in a demonstration school for culturally deprived children. Kennedy and Willcutt (1964) concluded from their review that ". . . when one corrects for practice, as with the use of a control group, praise is a reasonably stable incentive from study to study, contributing an incremental effect upon the performance and learning of school children [p. 331]."

Leith and Davis (1969) compared the effects of teachers' verbal comments with written comments on the performance of 13 year olds in a programmed course in logarithms. They found no difference between the effects of teachers' verbal and written comments. Post hoc analyses, however, revealed that negative verbal statements by the teacher (i.e., reproof) depressed performance significantly more than the teacher's neutral statements.

Another interesting effect of teacher comments is demonstrated by an experiment which coupled the dispensing of a food reward during snack time with verbal statements by the teacher (Risley, 1968b). The criterion measures of performance in those studies were both the number of verbal statements about what the children had done and the factuality of these statements. When the teacher gave the snack contingent upon a child's statement that he had, for example, played with the paints, the frequency with which the child made such a statement increased whether or not he had, in fact, actually played with the paints. On the other hand, when the teacher continued to give food contingent on verbal statements of certain play behavior, but added a comment

each time about the truthfulness of the response, performance of the child changed as follows. When the teacher confirmed the truthfulness of the child's statement by saying "that's right, you really did play with the paint," the rate of the child's reporting remained high. When the child's statement did not conform to fact, the teacher still gave him a snack but added, "you didn't really play with the paint though did you." The number of claims to have played with paint made by the latter group notably dropped. Thus, even though a material reward was contingent only on the child's verbal statement, the teacher's comment disconfirming the truthfulness of the claim resulted in reduction of untruthful statements.

Some studies have tried to identify differential effects of praise on students of differing characteristics and backgrounds. Two studies reviewed by Kennedy and Willcutt (1964) showed that fifth grade students labeled as introverts achieved a higher level of performance when praised, and, conversely, those students labeled as extroverts achieved a higher level when blamed. Leith and Davis (1969) also found that 13 year old students rated as both anxious and introverted performed better for praise than for negative incentives on a programmed logarithm course, a difference that was not found in the other personality types examined. The same researchers found no differential preference for praise vs. neutral or negative incentives relative to performance of students from high and low socio-economic home environments. Spence (1966) found both middle and lower class children performed better for verbal praise than for candy rewards.

Studies have also been conducted to identify characteristics of incentive givers that enhance the reinforcement value of praise; other studies have tried to alter the reinforcing power of praise givers. Unikel et al. (1969) found that female experimenters achieved better results than male experimenters with 5 and 6 year old White rural area children. Vega, in a study reviewed by Kennedy and Willcutt (1964), found that second, sixth, and tenth grade Black students improved their performance under "blame" conditions by Black examiners whereas, under White examiners, "blame" depressed scores of Black students. McArthur and Zigler (1969) tried to manipulate the attractiveness (valence) of the reward giver by showing one of two films to each subject before the latter performed a discrimination learning task. One film portrayed the experimenter as vituperative and the other film presented him as warm

and supportive. Those second grade boys who viewed the warm, supportive experimenter on film persisted longer on a boring, dull task in order to earn experimenter praise than did the boys who viewed the same experimenter portrayed as an erratic, abusive character.

In a study examining the effects of frequency of praise, Clark and Walberg (1969) found that more frequent praise conditions produced the greatest improvement in standardized reading test scores after a three week experimental period. Mech, Kapos, Hurst, and Auble in a study reviewed by Kennedy and Willcutt (1964) found that massed praise was more effective than spaced praise in enhancing the performance of fourth graders.

Cormier (1970) found that non-contingent praise compared to little or no praise was more effective in improving performance of eighth grade students from low income homes. Cormier postulated that, under non-contingent reward conditions, the teacher acquires positive descrimination cue properties. In other words, depending on the student's previous history of contingent rewards and punishment in school, even non-contingent rewarding behavior is a cue for productive behavior. In line with this hypothesis, McAllister et al. (1969) applied the following procedure with a lower track junior-senior high school English class. The teacher, a 23 year old female, was instructed to "disapprove of all instances of inappropriate talking behavior whenever they occured with a direct, verbal, sternly given reproof . . . using students names when correcting them." Also, regardless of classroom behavior, the teacher was to praise the whole class for being quiet. Praise was dispensed on a time schedule of fixed intervals specified in advance of the study. Compared with baseline, compared with a control class taught by the same teacher, and also compared with behaviors receiving the experimental treatment at other points in the sequence, this experimental procedure was very effective in reducing inappropriate talking.

Some experiments which focused on individual students demonstrated that, as the student's behavior changed under the conditions manipulated by the experimenter, other students in the class began to interact more with the target student and to augment the experimenter's contingency management program. Wasik et al. (1969), for example, reported their informal observation of class contingencies on the behavior of a target student three months after the experiment ended. "She [the target student] seemed to receive a lot of

attention from peers who both praised her when she was appropriate and warned her to change when she was not. She also spent more than the required time on her workbook [p. 193]." Other studies have directly manipulated the social pressures of the class on the behavior of individual students. Schmidt and Ulrich (1969) employed what Schwitzgebel (1970) might call a "behavioral prosthetic" device to control classroom noise. They used a sound level meter and a simple kitchen timer to monitor duration and intensity of noise in the classroom. Students earned recess time by maintaining noise level below 42 decibels for the length of time that the timer was set. A class of fourth graders was selected for the experiment because of its excessive noise during free-study period. The authors reported "peer consequences in the form of threatening gestures, arm moving, and facial expression . . . directed at more noisy members of the class [p. 174]." Packard (1970) employed another behavioral prosthetic device, a Cramer 1/100-second timer, to control attending behavior of kindergarten, third, fifth, and sixth grade classrooms in a small rural school. The teacher flipped the switch on the timer only when all students in the classroom were attending to desirable learning activities. All students earned rewards in proportion to the amount of time recorded in which all students were attending. A red light on the timer indicated that good behavior time was being recorded. In all classes the procedure improved the "whole class" attending behavior of students.

Knowledge of Results. A major assumption in programmed instruction has been that feedback of correct answers is reinforcing. A student was expected to be highly motivated to learn if he received immediate corrective feedback. In a study with 621 eighth grade students, More (1969) evaluated the effects of four intervals of delay in feedback on acquisition and retention. Students were given multiple-choice tests on one of two articles on science and social studies. All subjects read one of the articles, took $test_1$ on the article's content and then received the correct answers (IF) in one of four conditions: (1) immediate feedback after each item, (2) 2 1/2 hour delayed feedback, (3) one day delayed feedback, or (4) four day delayed feedback. Half the students took $test_2$ immediately after IF (acquisition), and the other half took $test_2$ three days after IF (retention). Immediate feedback yielded the poorest results for the acquisition group. Immediate feedback and four-day delayed feedback yielded the poorest results for the retention group.

Anderson (1967) reported a study by Melching in which feedback was given only when a student requested it. Such requests were made on only 31 percent of the frames; errors were made on 4 percent of the frames for which feedback was not requested and on 28 percent for which it was requested. Anderson did not report any retention effects or comparisons with groups receiving immediate feedback. Marshall (1969) compared various feedback conditions for 160 Caucasian kindergarten children performing a discrimination learning task on a marble game. He found that immediate verbal feedback produced more rapid learning than did accumulated points for toy or candy prizes when such knowledge of results was delayed by about six seconds. Several studies (e.g., Krumboltz & Weisman, 1962; Sullivan, Baker, & Schutz, 1967), however, have demonstrated that no knowledge of results produced learning rates as high as those produced by continuous feedback of correct results and that fixed and variable ratios of feedback hindered learning. Anderson (1970) has recently presented a convincing case for an attention-based interpretation of this type of result.

Feedback to teachers has also been used as an incentive device. Hall, Lund, and Jackson (1968), for example, daily informed each participating teacher about their students' progress by showing the teachers charts of the observed student behaviors. The authors reported that those ". . . daily contacts, plus weekly conferences in which the procedures were discussed and the teacher was praised for bringing about the desired behavioral changes, may have been central to the process of a successful study [p. 11]." This observation, however, was incidental to their study; it did not represent a systematic evaluation of feedback as an incentive to teachers. A study by Panyon, Boozer, and Morris (1970) did evaluate the effect of performance feedback to hospital attendants on the latter's use of operant techniques with patients. The names of attendants in four halls of the hospital were published each week along with the amount of time each attendant performed the desired task. Each attendant received his own feedback plus the feedback for all the other attendants. In all four halls the amount of attendant task performance increased sharply when public feedback was introduced.

Smith (1966) found that, in a classroom in which students were allowed to progress at their own rates, the interstudent competition implied in displaying charted progress before the class was a highly effective incentive.

Secondary Reinforcement. There is still lack of agreement among theorists about what is a primary reinforcer and what is a secondary reinforcer. In an effort to avoid the controversy, the term "secondary reinforcer" in this section is reserved for those incentives that clearly have no power to reduce any innate need within the organism. Their value is established by (1) their rate of exchange for goods, services, and privileges or, (2) their symbolic value for representing accomplishment. In this category would be included points, stars, tokens, money, trophies, etc.

Money was used as an incentive to improve reading accuracy and speed on the Gray Oral Reading Test (Cotler, 1969). The 96 subjects, who were fourth, fifth, and sixth grade boys, were each given 75 cents before the experimental session. Boys in the reward and punishment experimental condition were told they might win more money or lose some money depending on their reading performance. Control subjects were also given 75 cents before the experiment but told it was simply payment for participating. The contingent reward and punishment condition yielded results that were not significantly different from the control group results. The finding has been confirmed in other studies (e.g., Prince, 1967; and Smith, 1966). Alschuler (1969) went one step further and used "make-believe" money instead of real money. Each student in his study signed a contract with the teacher in which the student agreed to complete so much math for so much "make-believe" money. This so called "math game" resulted in significant improvement on standardized mathematics test scores over the period of a school year.

As has been mentioned earlier, public records of rewardable behavior can be used in a secondary reinforcement system. Circling numbers on cards (Clark & Walberg, 1969), writing names on the blackboard (Schmidt & Ulrich, 1969), placing marbles in a holder (Marshall, 1969), and writing out a little blue ticket (Surratt, Ulrich, & Hawkins, 1969) are just four of many creative examples presented in the literature.

The value of secondary reinforcers can be kept high by manipulating the exchange incentives. In some studies tokens were exchangeable for candy or toys (e.g., Heitzman, 1970; McMains, 1969; O'Leary, Becker, Evans, & Saudargas, 1969). In other studies they were exchanged for tickets to special events (e.g., Bushell, Wrobel, & Michaelis, 1968), play activities, or special privileges (e.g., Packard, 1970). In some situations, all candidates for

tokens may not be aware of the desirability of the exchange incentives, introducing a confounding variable into these situations. In order to control this variable and to enhance the value of tokens, Ayllon and Azrin (1968) employed a technique they called "reinforcer sampling" in which candidates were given brief exposure to the exchange incentives without any cost in tokens. They were, for example, allowed to view the first five minutes of a movie and then given the opportunity to view the rest of the movie for a "token" fee. This procedure is somewhat akin to that used by commercial airlines, which show a movie in full view of all the passengers then rent the earphones, which provide access to the movie's sound track, for $2.00 per set.

There are examples of investigations in which the reward, made contingent on performance but available at a later date, was mediated simply by verbal promise. One such example is found in an experiment by Halcomb and Blackwell (1969) in which ". . . performance of a monotonous, boring, and sometimes noxious task" was secured by promising credit toward successful completion of a college course. Those subjects for whom credit was made contingent on a certain level of performance quality achieved much higher scores than those subjects who were promised credit simply for participating. Thus, the verbal promise of reward also carried a certain motivational value which greatly influenced performance.

Aversive Incentives. Aversive stimuli receiving systematic study in the classroom have primarily been reproof, reprimands, disapproval, or simply indications that the student's response was wrong. Just as positive statements signify to the student which behaviors are desirable, punishment delivers information that the student's behavior is undesirable.

In some situations, however, punishment has increased the frequency of the very behavior it was supposed to decrease. In a functional analysis of punishment, for example, Becker and Armstrong (1968) found that tripling the number of disapproving remarks by the teacher increased disruptive behavior of primary students. "The teacher reported, 'When I stop praising the children and make only negative comments, they behave very nicely for three or four hours. However, by the middle of the afternoon the whole classroom is chaotic.'" [p. 43]

The complexity of punishment as a means of behavioral control is shown in a summary by Bandura (1969) of the factors that interact with punishment.

> In addition to the reward contingencies maintaining the punished behavior, the effects of punishment may vary considerably as a function of many other variables . . . including the intensity, duration, frequency, and distribution of aversive consequences, their temporal relation to behavior to be modified; the strength of punished responses; the availability of alternative behavioral patterns that are positively reinforceable; the presence of discriminative stimuli that signify the probability that a given performance will result in adverse consequences; the level of instigation to perform the negatively sanctioned behavior; and the characteristics of punishing agents. [p. 295]

This catalog of possibly related variables influencing punished behavior touches on three reasons which are often presented to explain why punishment increases undesirable responses. One explanation is that undesirable behavior often elicits teacher attention whereas desirable behavior does not. If teacher attention is rewarding to students, then undesirable behavior is functional to the students in winning that attention. Another explanation sometimes given is that reprimands and reproof are discriminative stimuli for undesirable behavior. A third explanation, one offered by O'Leary and Becker (1969) to explain their results, is that the teacher who "yells" at her students elicits undesirable conditioned emotional reactions which, in turn, provide stimuli for continued and perhaps more varied inappropriate behaviors.

In four experiments by Hall, Axelrod, Foundopoulos, Shellman, Campbell, and Cranston (in press), certain punishment procedures were found effective in reducing undesirable behavior. In one experiment the teacher pointed her finger and shouted, "No," whenever the seven-year-old deaf girl, who was the target subject, bit or pinched herself or another student. In another experiment the teacher decreased the whining and complaining behavior of a seven-year-old boy by contingent removal of slips of paper from the boy. Five colored slips of paper with the boy's name printed on each one were given to the boy at the beginning of each reading period. In two other experiments, out-of-seat behavior and grades of D and F on French quizzes were punished by requiring students to stay after school. Three senior high students were subjects of the latter experiment. In all four experiments, contingent punishment did decrease the frequency of the undesirable behavior.

One point that seems important in regard to classroom management is that punishment delivered without emotions such as anger, disgust, rage, etc.

seems to be more effective. In the study by McAllister et al. (1969), teacher disapproval delivered sternly but without threat of consequences and in conjunction with increased praise to the whole class, was judged to have been a contributing factor to the success of that management plan. The experiments by Hall et al. (in press) did not involve corporal harm or ridicule and were systematically administered whenever predetermined transgressions occurred.

Summary of Stimulus Manipulation. Studies employing the use of material incentives have included food, toys, and other items usually inexpensive but considered desirable to potential recipients. This pragmatic approach seems to have led naturally to the system of supplying an array of material incentives hoping thus to maximize the reinforcement value of the one chosen by the subject. Material incentives were less effective than expected in laboratory studies of discrimination learning. In those studies, the timing, placement, or mode of delivery may have been responsible for lack of success rather than the type of incentive used.

Praise is perhaps the easiest and most natural of social incentives to use in educational settings. Praise given at the time the specific desired behaviors are exhibited has proven effective in increasing those behaviors. Some students are less responsive than others to praise. Also, some persons are more effective as praise givers than are others. It has been shown, however, that the effectiveness of praise givers can be enhanced by increasing the frequency of praise and by pairing praise with the administration of other incentives that are effective. A very powerful social incentive is the mobilized social power of the peer group. Group social incentives appear to carry both the promise of acceptance and the threat of rejection.

Once a student becomes involved in a learning task, it is not universally true that external incentives will improve performance. Incentive administration, even feedback of correct results, under some conditions may interrupt and retard performance. The available data suggest that the student is the best judge of when information feedback should be given. Feedback to teachers about student performance may be an effective teacher incentive. One might expect, however, that the sources of such feedback data would need to be high on face validity and relevance to desired educational outcomes. Finally, public feedback of results seems to offer opportunities for competitive situations in which successful performance becomes an incentive in itself.

Secondary incentives seem to be as effective as material or social incentives in influencing behavior, once a system for their delivery and exchange has been established. The additional value of secondary reinforcers is their flexibility. They can be adapted in countless ways to be both convenient and effective. Their chief value relevant to other types of incentives is that many highly reinforcing events can be made available to students contingent on desirable performance through the use of a form of "currency" such as tokens.

Punishment is a complex and, at this time, a rather controversial area of psychological study. Most proponents of precision teaching methods (cf. Meacham & Wiesen, 1969) advocate the use of positive incentives rather than negative ones, due to the more predictable effects of positive reinforcement. In cases where punishment is used, it seems advisable to deliver it automatically and without emotional overtones.

Response Manipulation

The various incentives discussed so far under stimulus manipulation are of the type which one person does to or offers to another person. Another class of incentives are those which the target person does himself. In this review, these are called "response" type incentives. Implied in the concept of response "manipulation" are restrictions and controls on the behavior of the target person and an agreement about when and if certain responses will be permitted. Such conditions commonly exist or can easily be introduced in schools; therefore, response manipulation appears highly relevant to the topic of this review.

High vs. Low Probability Responses. Premack (1965) postulated and provided evidence to support his theory that high probability behaviors may serve as reinforcers for low probability behaviors. Homme (1966) reported several applications of this principle (commonly called the Premack Principle) in classroom settings. One such application was with high school dropouts or adolescents judged to be potential dropouts:

> . . . These adolescents had the behaviors usually associated with 'street kids.' They spoke a hip jargon, some of them were discovered to be carrying knives, some of them wore their sun glasses at all times, and so on. The low prob-

> ability behavior for these Ss was getting themselves through programmed instructional material in subjects like arithmetic and reading. Most of the high probability behaviors used to reinforce these were of the conventional sort: time for a break, coffee, smoke, coke, and so on. However, there were some surprises. For some of these Ss, going through a program in Russian proved to be a reliable high probability behavior.

When it was discovered that spending time on the programmed Russian courses was a high probability behavior, then permission to work on Russian was reserved as a reinforcing response for work done on arithmetic or reading.

High probability responses vary from setting to setting and from time to time. The list of responses used as reinforcers with third, fifth, and sixth graders in Packard's study (1970) included use of a private study booth, sitting next to a friend, use of a class typewriter, and being a "teacher assistant."

Self-Management of Responses. Response manipulation is especially important as a class of incentives because it leads one naturally to the question of self-managed learning behavior. A long range goal of incentive manipulation is either that learning will come to provide its own intrinsic reward and thus be self-motivating, or perhaps more realistically, students will learn how to motivate themselves to complete some of their boring but necessary learning tasks.

Relevant to these long range goals is research which illustrates the use of self-management as a learning incentive. In this vein, Campbell and Chapman (1967) compared two groups of fourth and fifth grade students on performance in a programmed geography course. Members of one group designated as learner-controlled (LC) were given the course objectives and learning materials but no study plan. The other group designated as program-controlled (PC) was given a study plan for the course. In the former group, teachers helped students formulate their own study plan. Of interest here is the greater incentive value of the LC program. Over the eight month period of the course, students in the LC group showed an increase both in liking for geography and in preference for self-directed instruction, increases significantly greater than for the PC group. It is also of interest, however, that the performance of both groups was about equivalent.

Some experiments have been conducted in which subjects rewarded themselves for achieving some performance standard. In an experiment with third and fourth grade children in summer camp, McMains (1969) recorded the behavior of his subjects playing a bowling game at a time when the children thought they were alone and unobserved. Some children in the experimental condition were given verbal instruction as to what score they should earn before they helped themselves to tokens exchangeable for high incentive toys. Other children observed an adult play the game and reward himself. A third group was given both the verbal and the modeled conditions and the fourth group (control) was given neither. The combined verbalization-model condition produced higher standards for self-reward than any other group. Verbalization and modeling had about equal effects and both produced greater performance increments than the control condition. Bandura and Perloff (1967) performed a similar experiment with 80 lower elementary level children from a lower middle-class area. This experiment also involved a game played supposedly without a monitor. Subjects either set their own standard or it was set for them by the experimenter; and they either reinforced themselves or were externally reinforced by machine. Control subjects received either non-contingent rewards or no rewards. Rewards again were tokens. Self-monitored and externally applied reinforcement were about equal in sustaining voluntary responses, and both were superior to control conditions. Unexpectedly most children imposed very high standards upon themselves such that their work-to-reinforcement ratio was highly unfavorable.

Another possible response incentive is self-determined goals. Kennedy (1968) compared the arithmetic performance of third and fourth graders in a low socio-economic area. Some students were told to do their best, others told to set their own goals for the week, still others were given goals set by the teacher, and a control group was given no goal setting instruction. Students given specific goals by the teacher or having self-set goals did better than students simply told to do their best. Various achievement levels of students were also compared but, unfortunately, achievement level was confounded with the teacher variable in the study.

Summary of Response Manipulation. The concept of high vs. low probability behaviors appears to have great utility for school settings in arranging

reinforcements for children. A high probability behavior for one student may be a low probability behavior for another. Thus, if one student likes math, the opportunity to study math may be employed as an incentive for the student to complete other learning activities. Yet another student might be given the opportunity to play a game or read a novel as incentive to complete other learning activities he finds less exciting. The subgroup of response incentives that might be classified as self-managed incentives are somewhat equivocal as incentives per se, yet they do represent desirable long-range goals of education. In other words, self-determined goals may not be the most effective incentive device for a particular student, yet self-management of rewards, self-determination of a study plan, and self-set goals are all laudible education objectives. And, for some students, increasing opportunities to map their own learning strategies may produce significant motivation for academic achievement.

Target Populations

Although most of the studies illustrating types of incentives have concerned students, it is not at all necessary to consider students the only target population for incentives. Anyone involved with student learning and responsible for student achievement is logically a potential target for incentives. Target groups, therefore, would include teacher aides, teachers, parents, teacher's supervisors and administrators. Neither do incentives need, necessarily, to be administered to individuals. They might be given more productively to an entire classroom of students, to a school, to a teacher's union, to a school district, to a PTA, and so on. One other target population for incentive delivery is a group of model students who exhibit behaviors which, if imitated, would lead to better academic performance by the imitators. Incentives to model students are referred to as vicarious reinforcement.

Individual Placement of Incentives

Students are logically the most relevant population to receive incentives, and the great bulk of relevant studies cited thus far have utilized student targeted incentives. All sub-groups of stimulus and response incentives, except vicarious reinforcement, have been applied to students and it would seem that all could possibly be productively applied as well to the other target groups.

Discussions about incentives to teachers have usually inferred that incentives would be money (salaries) or fringe benefits. A field experiment by Leno (1958) may inadvertently have identified another type of incentive to teachers. He carefully explained to his experimental and his control teachers that the criterion measure of the study was a test of arithmetic story problems. The teachers, furthermore, were instructed to tell their students, both experimental and control, "that they are engaged in a little experiment to see how much progress they can make in twenty days." The greatest gain occurred in one of the control groups, a gain of 2.9 years in 20 days of teaching compared to the next highest score of 1.8 years gain by one of the experimental groups. This result clearly was unanticipated by Leno. It suggests the possibility that the verbal instructions may have provided an incentive to the control teachers in the nature of a challenge to "beat" the experimental groups.

Feedback about student progress has been applied in studies involving extra work or change of behavior on the part of teachers and teachers' aides. Such research designs often employ classroom observation to obtain criterion data that may be shown daily to teachers. Packard (1970) reported that some teachers in a rural western Michigan school where he conducted his study were very reluctant to participate. "Hogwash" was one teacher's evaluation of the project and a typical reaction was, "I'll give it a try for a few days, but I'm telling you my heart's not in it." The rather dramatic changes in student behavior immediately recorded on the mechanical device induced the skeptics to give their full cooperation.

Information about the teacher's own behavior has sometimes been fed back to teachers as an incentive and as an instructional device. Researchers who observed and recorded student classroom behavior as part of their study often recorded teacher behavior at the same time and then used that data to

inform the teachers whether or not they were meeting the experimental conditions of the study (e.g., Hall, et al., 1968). Modeling of desired teacher behaviors has been employed in teacher training programs (e.g., Allen & Ryan, 1969; Lipe & Steen, 1970) in which vicarious reinforcement was systematically manipulated in conjunction with the modeled behavior.

Effective school administrators no doubt employ various incentives of both the stimulus and the response types such as selective praise at faculty meetings (vicarious social reinforcement), periodic written evaluation of teachers for personnel records (positive or aversive stimuli), and selective assignment of duties such as policing the lunch room, monitoring detention classes, and teaching high or low track classes (response manipulation). Some critics argue that often these incentives employed by school administrators are contingent on less than optimal educational objectives, such as order and conformity in the classroom. "This is particularly true in lower class neighborhoods," wrote Coleman (1969), "where order in the classroom is most problematic." [p. 22] Many of the behavior modification studies, in support of Coleman's assertion, have focused on behaviors that constitute classroom management problems, such as noise level (e.g., Schmidt & Ulrich, 1969), out of seat behavior (e.g., Osborne, 1969), and other disruptive behavior (e.g., Thomas, Becker, & Armstrong, 1968).

Hopefully the school is not the only place where academic learning occurs. Since homework is often formally integrated into student learning activities, it is appropriate to consider the home as a place for incentive manipulation. Hawkins and Sluyter (1970) involved parents as reinforcing agents; Ryback and Staats (1970) went one step further and involved parents as both tutors and reinforcing agents. In the latter study, parents of poor readers were given four hours of instruction in the Staats Motivation-Activating Reading Technique (SMART) and then were periodically supervised during the early weeks of a 5-7 month period that parents employed SMART to teach their children how to read.

It should be noted that parents were not the targets for incentive delivery in either the Hawkins and Sluyter (1970) or the Ryback and Staats (1970) studies. Rather, parents were simply involved in the process of delivering incentives to students in their homes. In one of the Head Start Programs (Herman & Adkins, 1970), parents themselves were paid a moderate

allowance for attending parent training workshops. That allowance was ostensibly given in order to cover transportation and baby-sitting costs of the parents. Singell and Yorden (1970) have proposed a system of payments to parents, half of which would be contingent upon parent attendance at various school functions and the other half of which would be contingent upon their children's achieving specific levels of academic performance.

Group Placement of Incentives

Several studies have demonstrated that incentives given to the whole class rather than to individual students could be effective in mobilizing class social pressure upon the behavior of its members. These studies were reported in the earlier section of this paper under the heading of social incentives (cf. McAllister, et al., 1969; Packard, 1970; Schmidt & Ulrich, 1969).

Coleman (1969) has suggested interscholastic "academic olympics" or other scholastic competitive games that could be structured on a group incentive basis. Zangwill's (1970) proposal for monetary grants to school districts based upon student achievement gains falls within this area. Also voucher systems such as those recently proposed by Jencks (1970) could be structured to provide group incentives to school districts.

Incentives to Model Students

A point often made by Bandura and effectively demonstrated in his research (1969, chapter 3) is that the observation of another person (designated as a model) being rewarded for a certain behavior can be a powerful incentive for the observer to perform the same behavior. It has been postulated that the observer reasons as follows, " If I do that, perhaps I'll get rewarded too." Thus, vicarious reinforcement would seem potentially to be particularly useful in inducing acquisition of novel responses that are prerequisite to improved academic achievement. Primary teachers achieve astounding feats of classroom management by rewarding individual students in view of the whole class. Statements such as, "See how quietly Johnny is sitting at his desk," seem to establish desirable models of classroom behavior which observing students then imitate. Flanders

(1968) reported in his review of research on imitation behavior that the combination of rewarding desirable and punishing undesirable modeled behavior has been found effective in inducing desired imitation behaviors in a large number of studies. Although directly relevant findings are not available, vicarious reinforcement would seem to have potential as an incentive device for academic achievement.

Summary of Target Populations

There are several relatively untapped target populations for potential individual or group incentive delivery. In addition to model students, these include teachers, paraprofessionals, parents, and administrators. A cost-effectiveness issue arises in regard to incentives delivered to parents or to groups of parents such as the PTA. Perhaps the greatest return on the incentive investment would be with regard to student behaviors that parents could monitor without any prerequisite special training. This, however, is an empirical question, not a conclusion. Another point with regard to the cost-benefit issue is that some types of incentives such as social incentives, knowledge of results, and aversive incentives are very inexpensive; some are even free. Their implementation, however, would require varying degrees of restructuring of the educational system. Some proposed voucher systems, for example, would result in relatively elaborate restructuring.

Modes of Incentive Delivery

Many ingenious systems of incentive delivery have been developed for or adapted to educational settings. The major objective of such systems is to link the incentive to the desired student performance. Some types of social incentives such as the teacher's smile or the teacher's proximity to a student usually require direct, personal delivery by the teacher at the time the student exhibits the desired performance. A few studies have employed mechanical devices to deliver incentives and other studies have introduced a system of withholding positive reinforcement as a means of extinguishing undesirable behavior. Still other studies have introduced various systems

designed to mediate between the performance of a desirable behavior and the delivery of an incentive when the latter occurs after a lengthy time interval. The ultimate goal, however, is for the student to learn self-discipline by rewarding himself for performing desirable, but perhaps tedious, study behavior.

Direct, Personal Delivery

Studies involving direct, personal delivery of incentives have usually been of the variety of teacher praise, attention, and approval contingent upon desirable student behavior. One variation of this pattern is found in studies by Hart and Risley (1968) and by Risley and Hart (1968) which paired teacher praise with the delivery of a snack to Black children enrolled in an experimental preschool. An incredibly simple but effective system of direct, personal incentive was used in the study by Benowitz and Busse (1970). In that study, four teachers of Black fourth grade ghetto children gave each child a box of crayons if he performed well on the Friday spelling test. In two other studies, disapproving statements were personally delivered by teachers to disruptive students in the class (McAllister, et al., 1968; Thomas, et al., 1968).

Teacher approval and disapproval are commonly accepted as appropriate and desirable incentives for the classroom. They are also easy to administer within the typical existing educational system. The direct delivery of approval and disapproval is, therefore, a logical starting place for incentive manipulation. Hall, et al., (1968) found that a beginning sixth grade teacher in a low socio-economic area was able to gain control over the disruptive behavior of her students by making praise and rebuke strictly contingent on students' desirable and undesirable behavior, respectively. This relatively easy incentive system, however, did not work for two other beginning teachers in the study. The latter did not achieve control over student behavior in their classes. Both teachers, who were unsuccessful in their use of direct, personal incentive delivery, changed to another type of incentive and mode of incentive delivery. Eventually they found a workable, although more elaborate, system.

Another arrangement for direct, personal delivery that permits the use of a variety of incentives is the individual adult-child conference. In a study involving second, fourth, and sixth grade students, Schwenn, Klausmeier, and Sorenson (1970) reported the results of teacher-child and teacher aide-child conferences held once a week for eight weeks. Over the eight-week period the adult employed vicarious reinforcement, approval, feedback, and goal setting as incentives to improve amount and level of reading and the reading skills of the conferees. Each child brought a record of the previous week's reading to each conference. That record provided the contingent basis for the various incentives. The conferences were effective in increasing the number of books read by both high and low achieving students at all three grade levels.

Mechanical Delivery Devices

Mechanical modes of incentive delivery have usually emitted auditory or visual signals when desirable or undesirable behaviors were being displayed. Perhaps the simplest such device is an answer sheet with a chemical means of changing color or other wise indicating that a correct answer has been marked. Another type of device was used in an experiment by Walker and Buckley (1968). A nine year old boy enrolled in an experimental class for behaviorally disordered children was given an auditory click signal to inform him that he had exhibited desirable study behavior and had earned another point toward a reward. A machine recorded the point at the same time the click sounded. Packard (1970) used a Cramer timer in a experiment to increase the amount of study behavior of kindergarten, third, fifth, and sixth grade children in a rural school. The teacher engaged the machine whenever she noted that all students in the class were studying and the machine, then, recorded that time interval. Whenever the machine was not recording time being spent in desirable behavior, a light visible to the students was automatically illuminated. In an experiment by Surratt, et al. (1969) the observer controlled a mechanical device that turned on a light at the desk of each target student separately when that student engaged in study behavior. At the end of each experimental session, the target students would gather around the recorder to see how much study time each one had accumulated and whether or not he had earned a reward. A similar device was used by Ward and Baker

(1968) to increase the amount of time spent in study behavior by first grade Black children. Interestingly, in the experiment by Surratt et al. (1969) a fifth grade student observed the target students for evidence of study behavior and operated the mechanical recording device. The fifth grade student was, in effect, the "behavioral engineer" of the first grade students in the study.

In an experiment by Schmidt and Ulrich (1969), a sound level meter monitored the noise level in a class of fourth graders in a school of mostly lower-middle class children. A human monitor recorded the variations in sound level over time and emitted a signal when the noise level went above a maximum tolerance level.

An electrical shock device was employed by Risley (1968a) to apply aversive stimuli to an autistic child when the child engaged in dangerous climbing behavior. Such aversive stimulus delivery devices have not been employed in classroom behavior modification studies although the paddle or a similar device has been known to be used in some schools.

Non-Delivery of Positive Incentives

Negative reinforcement may perhaps be classified as a mode of delivery, or rather, of non-delivery of incentives. Two types of non-delivery procedures most often used in behavior modification studies in the classroom have been the withholding of approval and the physical removal of the target person to a place of isolation, a "time-out" procedure. Negative reinforcement technically refers to the termination or withholding of aversive stimuli, but negative reinforcers have not been systematically studied as part of an incentive technique to improve academic performance.

Those studies in which the teacher was instructed to ignore disruptive behavior have typically included the instruction to also praise appropriate behavior (e.g., O'Leary & Becker, 1969; Wasik, et al., 1969). The assumption behind these directions is that the natural response of a teacher, generally is to attend to and then scold disruptive behavior when it occurs in the classroom. Quiet study behavior, on the other hand, does not naturally attract the attention of the teacher. If students are reinforced by teacher attention, then disruptive behavior is an effective way to gain that attention.

Thus, if the teacher withholds attention during disruptive behavior and simply ignores it, lack of reinforcement will "extinguish" the undesirable behavior.

Madsen, Becker, and Thomas (1968), in an experimental analysis of classroom behavior, independently varied rule giving, ignoring disruptive behavior, and praise of appropriate behavior. A kindergarten and a second grade teacher volunteered for the study; both classes were average to above average in intelligence. After a period of baseline measures, the teachers and their students devised rules for classroom behavior. The teachers carefully explained the rules and displayed them to their classes. The rules were explained again each day and displayed before the class, but other teacher behaviors were continued in whatever way was natural for each teacher. During the next time period of the study, in addition to repeating and displaying the rules, the teachers were instructed to ignore all disruptive behavior and all infractions of the rules. The following phase added yet another condition to those of the preceding phase. The teachers were told to continue the rules and to ignore all disorderly behavior as before but also to look for and express approval of appropriate and desirable student behavior. The authors concluded that rules alone had little effect, that the consequences of rules plus ignoring inappropriate behavior were not clear from the study, and that the combination of rules plus praise and ignoring was very effective.

A "time-out" procedure was evaluated by Wasik et al. (1969) as a means of controlling the behavior of two second grade Black girls in a demonstration school for the culturally deprived. There were two teachers in the classroom with the two students. Desirable and undesirable behaviors of the target girls were observed and recorded as dependent measures over the baseline period. Then the teachers were instructed in the following "time-out" procedure:

> Whenever a child's behavior met the criteria for social isolation, she was first verbally redirected to an appropriate behavior. If she continued to behave unacceptably, she was then warned: "You are not supposed to be doing . . . You are supposed to be doing . . . If you do not change, you will have to be taken to the quiet room." The teacher then turned from the child for 15 sec to allow her to make a change. At the end of the 15 sec, the child was taken to isolation only if she had not made any

> definite move to begin the structured activity and to terminate the unacceptable one . . . The child remained in the room for 5 min . . . . [p. 184-185]

In addition to this time-out procedure and otherwise ignoring inappropriate behavior, the teachers praised appropriate behavior. The procedures drastically lowered the frequency of undesired behavior.

Token Economies

Sometimes it is impractical or impossible to make a reward available at the exact time a desired behavior is exhibited. If a frequently occuring behavior of short duration is to be continuously rewarded when it occurs, a mode of delivery is needed that will not interrupt that behavior. For example, a teacher might want to reward a student's study behavior but not want to interrupt the task performance. In such cases a symbolic representation of the reward may be substituted in lieu of the actual reward delivery to be made later. Because the symbolic representation has often been in the form of washers, poker chips, or fake coins, the system of delivery has acquired the name of "token economy." It is, in fact, an economic system in which "earned" tokens may be exchanged, at a pre-established rate, for goods, services, or activities. Tokens have the advantage that the exchange rate can be varied; tokens can be given out rather unobtrusively; and tokens given at the time a desired behavior is demonstrated create a contingent relationship between the desired behavior and the substantive reward made available at a later time.

In the study by Lovitt and Curtiss (1969), students in a class for children with behavior disorders earned points for completing certain amounts of mathematics, reading, spelling, or writing. Bushell, Wrobel, and Michaelis (1968) conducted a study with summer session preschool children in which the children could earn tokens by attending quietly to instruction, working independently, or cooperating with other children. The tokens, in turn, could be used to purchase a ticket to a special event that took place near the end of each school day. A variation on the token system was introduced by Surratt et al. (1969). Four first grade pupils were told, "If you study a great deal during class today, you will get a little blue ticket. On that ticket you may write anything you want to do tomorrow morning." In the study by Clark and Walberg (1969), the participating teachers simply told individual middle-elementary level students to circle

numbers on a card to represent points toward a later reward. Hawkins and Sluyter (1970) reported the use of notes from the teacher to students which the students were to take home to their parents. The notes, in turn, informed the parents about the successes of their children. The parents, then, were instructed as to how they might enhance the "motivation" of their children by rewarding their children whenever such a note was brought home. In this way the notes served as tokens that were contingent upon specified levels of performance and which could be cashed in at home.

One of the most thoroughly researched token systems is the one developed by Staats and his associates which came to be identified as the Staats Motivation-Activating Reading Technique (SMART). The technique has been effectively used by parents, high school seniors, and adult volunteers from the P.T.A. to improve the reading of students, age 8 to 16, who had exhibited a reading deficit (Staats, Minke, Goodwin, & Landeen, 1967; Ryback & Staats, 1970). Tokens of different colors were given specific money values and were dispensed for accurate reading responses and scores on tests. If the person administering the Staats' technique refrained from expressing negative comments or gestures, the technique was successful in sustaining learning trials over 50 hours and individual word responses approaching 100,000. The achievement measure that showed the most consistent improvement was the number of new words learned and retained.

There is some controversy about how to maximize the reward value of tokens. On the one hand, Ayllon and Azrin (1968) discuss reward sampling by mental patients so that patients are fully aware of what they can buy with their tokens. Those authors presented evidence that even patients who had previously purchased a certain reward were more likely to purchase that reward again following a brief sampling. Stuart (1970a), on the other hand, suggests that knowing the reward in advance may tend to devalue it. Thus, Bushell _et al_. (1968) may have heightened the reward value of the special event ticket by not announcing ahead of time what the special event would be each day.

Performance Contracts

In some incentive delivery systems a formal or informal contract is made between the incentive giver and the incentive receivers. The contract may take the form of a brief statement by the incentive giver specifying the

rewards that will accrue for certain behaviors (e.g., Benowitz & Busse, 1970; Osborne, 1969; Halcomb & Blackwell, 1969). Other contracts may be written out and formally agreed upon by both parties (e.g., Homme, 1966; Cantrell, Cantrell, Huddleston, & Wooldridge, 1969; Smith, 1966). Interestingly, the practice of performance contracting in the classroom between teacher and students has more recently been applied to formal and binding arrangements between school districts and private firms (Education Turnkey Systems, 1970).

A contract may be written by the incentive giver or it may be drawn up by both parties together. The experiment by Lovitt and Curtiss (1969) tested the performance consequences of contracts in which reward contingencies were specified by the teacher. The recipient was a 12 year old member of a class for children with behavior disorders. Points were earned for work satisfactorily completed in all academic areas, and they could be redeemed for time spent in a "high interest" room. The high interest room contained many possible activities that were desired and valued by the subject. The authors found that more academic work of a given quality was performed when the student specified contingencies than when the teacher specified them. Although the student rewarded himself at a higher rate than the teacher usually did, when the teacher specified the student reward contingencies at the same high rate, student rate of performance was not as high as it had been under the student-specified contingencies.

Reinforcement Menu

Another mode of delivery has to do with the arrangements of rewards. A behavioral engineer usually wants the incentive to be maximally rewarding and easily and inexpensively administered. This objective is complicated by the fact that an individual's tastes and interests fluctuate from day to day. A person may get tired of one type of reward after receiving it several times and he may not be willing, then, to work for that reward.

In an attempt to impose some reason upon this fickle side of human desires, sometimes called whim, Helson (1964) postulated that the typical level of stimulation for a particular individual may be labeled the adaptation level and that a certain deviation from that optimum level, either upwards or downwards, will be maximally reinforcing to that individual. Berlyne (1969) postulated that such qualities as novelty, incongruity, suddenness, and intensity of stimuli could arouse curiosity and subsequent interest

in that set of stimuli. All of these qualities, however, are dynamic, requiring continuous creative effort in order to maintain their curiosity-arousing properties.

One solution to the problem of incentive satiation is to provide not one but an array of incentives from which the incentive receiver may choose that which appeals to him most at a particular time. For students, a separate room or special area is sometimes reserved and stocked with a large variety of games and activities. The potential population of incentive receivers is perhaps the best source to be polled as to what should be included in such a room.

In the study by McIntire, Davis and Pumroy (1970) a special project room was divided into three areas marked off by identifying colors and distinguished by the levels of attractiveness of articles and activities within each area. The most attractive activities were enclosed in red. A smaller section was marked off by yellow; the smallest section, and least interesting, was white. To get into the red area, the child had to get either a 90% or better (points correct over total possible points) or go up 10 percentage points from his score of last session. To be allowed into the white level, the child could not go down more than 10 percentage points from his last score. Yellow was in between. Points were determined by assigned work in the classroom.

An array of possible rewards might include response-type incentives like those in the Packard (1970) study. Response-type rewards have been known to include the chance to work a specified period of time for the school's janitor (Meyerson, 1970, personal communication). Yet another type of incentive choice system is to award a certain amount of "free time" to the incentive receiver (e.g., Osborne, 1969).

Self-Delivery

It is not at all far fetched to conceive of a student as delivering rewards to himself for work completed. For those high school graduates who go on to college, the transition might be considerably less difficult if they learn in high school how to change their own behavior and personal habits. This would be true, for that matter, of anyone going into a situation of greater responsibility for one's own actions.

Winter, Griffith, and Kolb (1968) compared college students who were successful in engineering personal change with those who were not successful. They concluded that "the establishment of a goal is crucial in arousing motivation for the difficult struggle to achieve a change."

Fox (1962) counseled college students to help them learn how to manage their study behavior. One technique in the self-management plan was for a student to set a specific time each day for study and always to follow, never precede, that study time with a desirable social behavior such as having coffee with a friend.

It is clear that the learning of mathematics, reading, or any other subject requires some tedious, repetitious drill (Staats, 1968). A student is much more likely to succeed at such a task if he can create a game out of the task, space drill sessions so that they always precede more interesting activities, or set intermediate goals and compete with himself to attain those goals.

Summary of the Modes of Incentive Delivery

Since modes of delivery should be as effective as possible with the least amount of expense or inconvenience, the iterative approach to designing such a mode of delivery appears promising. As was done in Hall's et al. (1968) study, the first iteration would consist simply of delivering praise and reproof carefully contingent upon desired and undesired behavior. If that doesn't achieve the desired behavior, then some alteration of the procedure may be instituted and checked for its effects. The requirement for an iterative approach to incentive system development is a measurement technique for assessing the amount of behavior change on the part of the incentive recipients.

Mechanical devices in the classroom have been used primarily to inform students when they were performing behaviors that would receive a reward. Such machines, however, usually require a human to monitor the behavior and control the sending and withholding of the machine's signal. An exception to the latter was the use of a machine to actually monitor the noise behavior of students. Thus machines have been introduced for their constant vigilance, added convenience, lack of bias, and relatively unobtrusive incentive deliveries.

One of the most effective modes of incentive delivery for eliminating undesirable student behavior, especially for controlling disruptive behavior, is the withholding of attention or approval. If removing the teacher's attention to undesirable student behavior does not cause the behavior to extinguish, then perhaps it is the attention from other students that is maintaining the disruptive behavior. In the latter case, "time-out" is an alternative approach that appears to have desirable results.

A "token economy" is a system for mediating between the desired performance and a reward delivered at a later time. What appears to occur, however, is that the tokens, themselves, acquire secondary reinforcement value. Thus, in phase I of the engineered classroom (Hewett, Taylor, & Artuso, 1969), students exchanged tokens for tangible rewards whereas in phase III, the tokens were simply exchanged for a mark on a graph showing the number of tokens earned. In addition to its convenience and flexibility, then, a token economy is usually a very effective system of incentive delivery.

The reinforcement menu has been offered as a solution to the problem of reinforcement satiation. It also provides a system for graduated rewards and withholding the more desirable rewards for superior performance.

Finally, self-delivery is presented as the ultimate goal, one that does not rely entirely on the intrinsic motivation of learning materials to maintain student volition in learning.

Timing of Reinforcements

The fact that the timing of reinforcement influences behavior is illustrated by comparing a student's study behavior in two separate courses. In one course only a final examination is given and in the other course weekly tests are given. Student study behavior, engaged in as a means of preventing a poor grade, would probably look very different in the two courses. The frequency of failure-avoidance study behavior in one course would likely show a dramatic increase near the end of the semester, whereas in the other course it would show weekly increases.

Two dimensions of incentive timing or scheduling will be discussed here. Immediate vs. delayed is one dimension. The other concerns continuous vs.

fixed vs. variable interval delivery. The point to be emphasized is that the experimental analysis of behavior has shown that differential behavioral patterns are produced and maintained by each of the different schedules of reinforcement (Ferster & Skinner, 1957).

Immediate vs. Delayed Delivery

One of the major generalizations of behavior modification research is that incentives must be made contingent upon the behavior they are to increase or decrease, otherwise they will have no effect. The weight of findings cited thus far shows that careful timing of praise and reproof so that they immediately follow and are made contingent upon desirable and undesirable behaviors, respectively, produce dramatic changes in student behavior. In the study by Reynolds and Risley (1968), social praise and attention made contingent on talking behavior of the target child significantly increased talking behavior. But, when the same incentives were used to differentially reward other behaviors, then the child's verbalizations decreased.

As pointed out earlier, the contingency requirement for incentives to be effective is more difficult to fulfill when reward cannot be given immediately or when it comes at the end of a long chain of behaviors. In the study by Benowitz and Busse (1970), a good score on the Friday spelling test could earn a box of crayons; but that score could not be earned by most children unless they studied the words every day of that week. The teachers mediated the connection between daily behaviors and the delayed reward by daily announcing the possiblity of a reward to come on Friday. Setting goals and signing contractual agreements are other ways of establishing a connection between a chain of unrewarded behavior and terminal rewarded behavior.

Continuous vs. Fixed Interval (Ratio) vs. Variable Interval (Ratio) Delivery

The many types of reinforcement schedules applied in the experimental analysis of behavior are not practical in most field situations (an exception may be programmed instruction). If the monitor is a human, in most classroom situations the teacher, then a rigid schedule of surveillance and reward cannot usually be followed.

The usual pattern is to begin by frequently rewarding desired but relatively new and unestablished behavior. Prince (1967) began his study by continuously reinforcing every correct response with a penny. Then he gradually shifted to a variable-ratio schedule until the average frequency of correct response was nine for each reinforcer. After shifting to a variable ratio, in some studies, the ratio continues to be decreased gradually until reinforcers are phased out. In the study by Schmidt and Ulrich (1969), a timer was set at variable intervals. Every time the timer bell rang, any student found out of his seat had to forfeit five minutes of his gym time. The average length of intervals between bells was gradually lengthened until finally they didn't occur at all.

Procedures designed to eliminate undesirable behavior often are naturally phased out. "Time-out" procedures, for example, are usually very effective in that the behaviors on which "time-out" is contingent very quickly terminate. Avoidance training of this type is somewhat akin to that of Solomon's shock avoidance training of a dog (Solomon & Brush, 1956). The animal must respond before the end of the few second interval in order to avoid a painful shock. Unless the dog periodically delays responding for more than the established interval, he does not learn whether or not the shock has been terminated. Remarkably long periods of responding have resulted from such techniques.

CRITERION MEASURES FOR INCENTIVES DELIVERY

Included in this chapter are suggestions for establishing educational goals for students which would be suitable as targets for incentive delivery. Closely related to educational goals are the techniques and procedures utilized to measure their attainment; in fact, many educational goals may be defined by their measurement operations. Therefore this chapter also looks at the advantages and disadvantages of available techniques for arranging incentive delivery based upon measures of student performance. Measurement procedures are suggested which would closely fit the needs of an experiment investigating the effects of educational incentives. In general, it is proposed that these procedures would include standardized norm-referenced achievement tests, criterion-referenced measurements, and behavioral observations.

Educational Goals

Clearly one of the highest national priorities is that all American youth should acquire basic reading and computational skills. These basic skills are generally taught in the elementary grades; however, a predictable percentage of children, who generally exhibit certain low socio-economic characteristics, do not acquire these skills in the early years. This subpopulation drops further behind its age-grade peer group, and although its members often seem to acquire through practical experience a number of necessary reading and computation capabilities, their orientation to formal teaching activities becomes less and less pronounced. As a result of this situation, they may come to exhibit a behavioral syndrome summarized by the label "dropout." This syndrome, comprised of various maladaptive and disruptive sequences, is well known to the psychologist who deals with frustrated organisms. It has been suggested that a continuing and deepening failure cycle of this nature may best be combated by interrupting it at its origin; namely, by developing diagnostic and compensatory programs which are successful in teaching the basic reading and computational skills to children who have

neither those behaviors nor a middle class reinforcement history. Although much remains to be learned about the optimal form of such programs, the present authors feel that enough is currently known about successful techniques to offer improvements over existing programs (see Hawkridge et al, 1968, 1969; Messick, 1969; Ulrich, Louisell, & Wolfe, 1970). As described in the Fourth Annual Report of the National Advisory Council on the Education of Disadvantaged Children (1969, p. 23), these include: (1) program of more than two months' duration; (2) clear statement, in advance, of program objectives; (3) instruction and materials closely relevant to the objectives; (4) high degree of individualization; (5) teacher training in the methods of the program; and (6) active parental involvement.

A preliminary examination of many of the most successful educational products currently in use reveals heavy emphasis on these components (for example, Sullivan, IPI, and DISTAR materials). Of interest here is the fact that sets of well specified objectives, stated in student performance terms, have been constructed for almost all these products such that their attainment may be measured objectively. In general, these objectives are arranged in such a manner that their sequential mastery defines an instruction strategy. This allows for diagnostic evaluation which can suggest both a student's current level of skill attainment and the skills he has yet to attain in order to reach a specified criterion level. It is suggested that in all experimental investigations of incentive effects, a concerted effort be made to establish sets of specific objectives for target students. Following are three alternative strategies which have been used to define criterion levels, along with a suggested resolution suitable for a large-scale incentives experiment.

Specific Behaviors

Although the use of incentives predates the carrot and oxcart, it appears that much of the current interest in student incentive techniques stems from recent educational applications of operant conditioning and the experimental analysis of behavior. The preceding review of literature on educational uses of incentives has demonstrated a preponderence of references from this source. Most of these studies exemplify the "experimental analysis" approach, stressing the use of a few subjects, keying on a few specific behaviors, and calling for fairly strict control of the environment over a prolonged period (Schutz & Baker, 1968; Sidman, 1960). Indeed, one of the primary

tenets of "precision teaching" (e.g., Meacham & Wiesen, 1969) is that the teacher should focus on very specific behavioral goals for one or two students at a time. Mastery of behavior observation and tabulation techniques is a prerequisite for teaching under these new methods. Only by using behavioral observation techniques can reinforcement delivery truly be made immediately contingent on various behaviors; this is a necessity for shaping of behavior and scheduling of reinforcement.

Several generalizations can be made about the requirements for specific behaviors as criteria for incentive delivery. The first, of course, is that they must be public events open to reliable observation. Behavioral data of this type are tabulated in frequencies per unit of time and have proven to offer an attractive and powerful alternative to other techniques for collecting observations of human behavior (for example, the testing paradigm, cf. Skinner, 1966b).

The second generalization is that the behaviors should be important; they should have a demonstrated relationship to the attainment of various goals of education, since rewarded behaviors will likely increase to the exclusion of other behaviors that might also be desirable. It is easier to observe and tabulate the frequency of various types of disruptive behaviors, studying and attending behaviors, etc., than "academic achievement" per se. Needed in this case is a clear and empirically demonstrated rationale relating such behaviors to measured achievement. Davidoff (1970) and Cobb (1970) present research which appears to be directed at this problem. The approach of the latter author, in particular, seems promising. After having demonstrated and cross-validated a relationship between selected behavioral categories (such as attending, volunteering, compliance, peer-to-peer conversations, etc.) and scores on the Stanford Achievement Test, he has initiated procedures aimed at increasing test scores by increasing the frequency of positively related behaviors.

The final generalization is that the observations must be practical. Large sums can be spent in obtaining reliable observations of classroom behavior, using methods such as videotaping; furthermore, it is sometimes difficult to obtain immediate feedback from the results of such observations. At the other extreme, the teacher is often too busy with instructional activities to engage in prolonged student observation. Systems utilizing

the services of a paraprofessional observer may offer at least a partial solution to this problem. Preliminary data from such systems (e.g., Petrafeso, 1970; Bushell, 1970) seem encouraging.

Standardized Tests

Standardized achievement test scores are widely used and interpreted freely by educators throughout the country. Well conceived testing programs are becoming the rule rather than the exception in many of the nation's larger school districts (Badel & Larsen, 1970). Since achievement is often defined in terms of gains on standardized test scores, it is understandable to find these tests adapted freely for the purpose of measuring student performance as a criterion for accountability. In contrast to classroom observations of behavior, standardized test scores are easy and convenient to collect. In addition, such scores enter nicely into formulae which can be agreed upon well in advance for calculating payments to performance contractors; there is little ambiguity or room for subjective judgment at collection time (disregarding, of course, the problem of coaching). Payment formulae typically result in the payment of a bonus for each student who makes a specified grade-equivalent gain; many formulae also result in the imposition of a penalty for gains below a certain level. Unless major changes come about in the prevailing practices, there is little reason to believe that large-scale incentive application projects of the future can avoid the use of standardized norm-referenced achievement tests.

Several problems need to be considered in evaluating the potential of standardized tests as criteria for incentive delivery. First is the problem of immediacy. Long delays must be acceptable between the taking of the test and the return of results. Such latency destroys any practical use of pure operant-type incentive techniques.

The most visible problem at present has been that of item coaching or "teaching for the test." In the Texarkana project on performance contracting, delivery of incentives to an outside contractor, Dorsett Educational Systems, was based upon grade level increases in student performance on the Iowa Test of Basic Skills and the SRA Achievement Series. Preliminary information (Education USA, September 7, 1970) indicates that from 30 to 100 percent of the items in the May posttests were directly taught in the Dorsett Rapid

Learning Centers during the course of the school year. This issue goes beyond the present problems of Texarkana to a broader question of educational program evaluation. If the achievement test is a proper measure of the educational results of a program, it should contain items which sample all of the important program objectives. On the other hand, if the educational program is to do its job, it must contain instructional activities keyed to the same objectives. Some overlap of content seems unavoidable; indeed, it is essential. Since the test scores themselves determine the contingencies of payment, the security of standardized test items may become the most important concern of involved educators. The maintenance of such security during the development and standardization of an achievement test is difficult at best, and after publication, it is impossible; moreover, it would be impossible to produce a new form of a standardized test to evaluate every contractual agreement.

Another problem involves the goals of various programs. Standardized norm-referenced tests are likely to measure a somewhat different set of objectives than those of a specific educational program. For example, a score on a published mathematics test may represent student performance on ten objectives; however, a given mathematics program may be concerned with only four objectives, and just two of these may be included in the standardized test. Thus the single composite score yielded by the test would not be appropriate for measuring the success of the mathematics program in terms of meeting its objectives.

From the psychometric point of view, much has been written about the dangers of using standardized test scores to evaluate the progress of individuals. The use of raw gain scores, or gains in grade equivalents, has continually been eschewed (Harris, 1963; Cronbach & Furby, 1970) because of the technical problems of reliability involved. Yet these problems are generally overlooked in establishing criteria for large scale incentive delivery. Needed are specific techniques for surmounting these statistical problems.

A final problem involves the fact that standardized tests are generally constructed to fit the capabilities of the average student at the given age/grade level of the test. Students who have a history of educational failure may find tests which are appropriate to their age/grade level generally

too sophisticated. It seems more feasible to attempt to administer standardized tests appropriate to the functional level of a student rather than to his age or grade. For example, it may be necessary to give a second grade level reading achievement test to a ninth grade student who is deficient in reading skills.

Criterion-Referenced Tests

Perhaps the most promising development in terms of establishing criteria for incentive delivery systems involves tests whose items are derived directly from some well-specified standard of performance. Glaser (1963) was among the first to point out the distinction between these tests, which reflect what the student can do, and tests which show how the student compares with others. Carver (1970) has presented the case for using such tests in the measurement of individual performance gains resulting from educational programs. When items are constructed so as to directly measure the degree of attainment of various behavioral goals of a program, an interesting marriage is achieved between constant behavioral observation and sporadic evaluation with norm-referenced achievement tests. Frequency of observation may be achieved along with ease and convenience of obtaining feedback. Such measurement is often connected with techniques for individualizing instruction, so that mastery of required concepts can be certified before a student moves along in his educational program (e.g., Lindvall & Cox, 1970; Shanner, 1968). Small increments in behavior can be detected by periodic small scale tests, enabling more frequent opportunities for incentive delivery.

The criterion-referenced test also offers advantages with regard to the coaching problem. Extensive tryout of items and standardization are not required, since the test acquires its validity primarily in terms of its relationship to the behaviors delimited by the criteria. Constant generation of a parallel-item pool whose members represent the entire set of objectives for a course could practically eliminate efforts to "teach to the test."

Klein (1970) has pointed out a problem in utilizing criterion-referenced measurement to make judgments about program effectiveness.

> To illustrate this point, let us suppose that a new course unit in 10th grade biology led to 30% of the students attaining all of the unit's 20 objectives, 50% of the students attaining 15 objectives, and only

> 20% of the students achieving less than 10 objectives. These results look very impressive and a school official might be very pleased with the effectiveness of the program. But would he still be happy if he discovered that most students could achieve 10 of these objectives before taking the unit, or that the criterion of attainment was 1 out of 5 items correct per objective, or that items used to measure an objective were not truly representative of the range of items that might have been employed, or that 80% of the students at other schools (having students of comparable ability) attained all 20 objectives using a criterion of 4 out of 5 items correct per objective? One expects that the school official would make a rather different evaluative decision regarding the program's worth had this latter information been available to him. Clearly, grade norms or other kinds of normative based data would help clarify the actual utility and significance of the program in achieving its objectives. [p. 3]

It will be noted that much of the strength of the criterion-referenced tests depends upon the specification of an adequate set of objectives from which items may be sampled. Other critical factors involve the construction of items which adequately represent attainment of the objectives.

A Resolution

In carrying out an experiment on educational uses of incentives, it will be necessary to identify in advance a set of objectives appropriate to the unmet needs of the selected student populations. In general, these will be cognitive domain objectives representing stages in the attainment of basic reading and computational skills; however, other objectives should be accepted provided that a clear rationale can be made relating them to the above skills. It may be possible to establish a set of objectives common to all educational programs. This would facilitate comparative conclusions to be drawn later. In reading, objectives could include aspects of readiness, sound-symbol relationships, word attack skills, and comprehension; in math, aspects of number recognition, sets, combinations, arithmetic operations, etc., could be included. Sources for such objectives include National Assessment (Tyler, 1966); PROBE and IOX-Instructional Objectives Exchange (Popham, 1970); the master objectives for Project PLAN (Flanagan, Shanner, & Mager, 1971) and the actual participants in an incentives experiment. Devices

for the measurement of attainment of these objectives will involve a combination of the three types of procedures outlined previously.

In cases where behavioral objectives of sufficient importance are identified, pre and post instruction observations of student behavior may be undertaken.

Standardized tests, when utilized, would be scored so as to yield data related to these objectives in addition to the normed composite scores, utilizing a method described by Cox and Sterrett (1970). When grade-equivalent gains are obtained, reference can then also be made to pre and post instruction performance on items actually related to the objectives of the program in determining the terms of an incentives contract.

Criterion-referenced tests would be utilized when available (or should be contructed) to produce similar data on degree of goal attainment. Such tests should be especially useful in determining pre-existing performance levels which could serve as a baseline for contingency contracts. An effort should be made in all cases to secure scores on a comparable reference group so that changes can be put into a meaningful perspective. As Carver (1970) has suggested, the most meaningful, useful, and simple statistic would be the percent of respondents that changed their performance on a given set of items over the course of an experiment. It would appear most productive to use this measure as the ultimate criterion for incentive delivery.

Regardless of the criterion of gain which is selected for determining incentive delivery, this gain should be considered relative to the previous measured ability and performance of the subject population. To expect the same rate of gain for a low ability group as for a high ability group is to go against predictions based on both past performance and common sense. This is definitely not intended to suggest that inferior achievement from disadvantaged or low socio-economic populations is to be expected as a matter of course. Indeed, academic <u>excellence</u> should be sought (Clark, 1970). However, given the same intensity of educational <u>treatment</u>, high ability students will probably gain more in terms of measured achievement.

MONITORING INSTRUCTIONAL PROCESSES

Incentives alone may provide motivation but they do not provide the means for improving student achievement. Nevertheless, during the interim between the time that incentive inducements are offered and the subsequent point in time when the target students' achievement outcomes are measured, something is expected to occur that will improve the target students' achievement. If greater than expected gains in student achievement occur, it is important to learn what the incentive receivers did to facilitate that gain. Specifically, it would be important to learn to what extent the incentive receiver applied some form of pressure on students and/or to what extent innovative instructional techniques and equipment were employed.

The intervening events between the incentive inducements and the student achievement outcomes could be identified simply by recording all those events. It would be more practical, however, to process only those particular intervening events that link incentives to student achievement and to obtain a quantitative measure of each event. With such quantitative data it would be possible to compare the tactics employed under the incentive treatment condition with the tactics used under the normal, existing incentive conditions.

In order to obtain such quantitative data, it is necessary, beforehand, to specify the categories of events that are to be measured. There are bases for hypothesizing about what tactics might be employed to enhance student achievement. We know, for example, what innovative materials and techniques are commercially or otherwise available. Manuals have been published which purport to instruct tutors how to motivate their tutees. Published studies of classroom observations suggest other means that might be used to enhance student achievement. From these and other sources, relevant tactics can be hypothesized and formulated into operational terms.

Even the most carefully devised operational categories might not provide for all of the events that are critical to the purpose of the study. Important data may be missed or some surprises may occur in the form of events that do not fit the measuring instrument. In order to minimize these pitfalls, the measurement techniques should be tried out in advance of the study. A pilot test would reveal potentially important tactics

that were not anticipated; it would also test the training procedures for the data gatherers and the reliability of the instruments.

Categorization of Tactics that Might be Used to Improve Achievement

There are no doubt several ways to conceptualize the tactics that might be employed by incentive recipients to improve performance of target students. Proposed here as a preliminary conceptualization are three major categories: innovative techniques, positive and negative incentives, and selective instruction.

Innovative Techniques. The recent explosion of innovations in education is awesome. Fortunately the task of identifying innovative techniques and devising methods for assessing their use is now being completed at the American Institutes for Research for use in a longitudinal study of demonstration educational programs. Some of the innovations of concern in that longitudinal study would not be relevant here because they affect an entire school or school district. Others, which affect individual classrooms, would be of interest. A teacher, for example, would not likely bring into his classroom an entire individualized system such as PLAN or IPI, but he might bring in programmed materials such as Sullivan. Those materials might then be employed in a quasi-individualization of instruction.

In addition to using materials suitable for the individualization of instruction, the teacher might introduce other innovations, such as team teaching, multimedia presentations, independent study, volunteer (or paid) aides, and home visitations. The types of innovations that might be employed by peer tutors or parents, of course, would not be as extensive as the potential innovations a teacher might use.

Innovations, by definition, are elusive because any creative change technically could be included. The incentive receiver might generate his own innovations rather than use someone else's. A taxonomy of educational innovations for purposes of this study, therefore, will need to be open ended to permit the recording of innovations that may not have a comfortable resting place in the existing jargon.

Positive and Negative Incentives. It seems possible that participating teachers would utilize incentive concepts in thier own interactions with target students. Nevertheless, even if incentive recipients did not design special incentive systems for their target students, they would certainly employ some types of positive and negative incentives,

simply because incentives are so common and so basic to the guidance of human behavior. It will be important to learn whether incentive recipients develop formalized incentive systems for their target students and to what extent they employ unstructured positive and negative incentives. The whole range of types of incentives presented in the literature section of this report will be considered as possible tactics that incentive recipients might use in order to enhance the performance of their target students.

Additional information about the targets of praise and punishment, or other incentives, would also seem to be very useful. For example, one might record whether incentives were contingent upon the student's ideas, expressed feelings, work practices, work products, or conformity to rules of behavior.

Selective Instruction. A third category of tactics to enhance achievement falls within the cognitive area, related to the structuring of instructional stimuli. It is conceivable that incentive recipients might focus their instructional efforts rather restrictively on the information usually assessed by selected measurement techniques. Other objectives pertinent to language and mathematics, such as the refinement of personal values, creative self-expression, and process skills, might be given short-shrift in the zeal of incentive recipients to promote pay-off behaviors.

Simon and Boyer (1967) presented a generalizable category system that may be applicable to gathering information in this area. Their system included data recall, data processing, and evaluation. Whereas data recall would correspond to simple rote learning, data or information processing might involve classification, definition, comparison, generalization, and so on.

Student Responses

A significant link in the chain of events between incentive inducement and student achievement outcomes is the immediate and ongoing response of target students to the incentive recipients' tactics. Teacher "A" may be an effective team teacher but Teacher "B" may not. Peer tutor "A" may employ positive incentives effectively to maintain target students on an

intensive learning schedule, whereas peer tutor "B" may not. It is proposed, therefore, that the quality of tactic implementation be assessed by measuring student responses to these tactics.

Within the time and budget constraints of the proposed experiments, it probably would not be possible to relate student responses to individual tactics of incentive recipients. It would be possible to assess the overall response of students to their environmental contingencies. One very important student response variable, for example, would be the amount of time spent in productive study versus time spent in non-productive or in transitional behaviors. A daily or perhaps weekly record of target students' attendance, completed assignments, and test scores would also be useful.

Measurement Tactics for Process Evaluation

Teachers and students often are burdened by many requests to provide data for research. Teacher and student contributions that take time away from the business of education must therefore be minimized. Existing records should be used whenever possible. Some data forms, however, would have to be completed by students, teachers, and parents.

There is no effective substitute for actual classroom observation of ongoing behavior as a source of information about many of the incentive recipients' tactics and students' responses. There are, however, many problems with classroom observation. It is not know, for example, how outside observers in a classroom might influence teacher and student behavior and thus contaminate the data. Training classroom observers must sometimes be extended in order to achieve a respectable degree of observer reliability; training can become expensive. If a new observation instrument has to be constructed, a period of development is required to test and revise the instrument so that it has discrete categories of behavior that can be reliably tallied by observers.

The measurement techniques for an incentives experiment must provide data in a form that will permit between-group comparisons. A second objective would be to provide generalizable knowledge about cause and effect relationships between tactics of incentive recipients and student performance.

Following are some examples of the types of procedures which could be utilized to collect existing records, self-reports, and classroom observations. Self-reports would be obtained from the incentive recipients and from the target students. Observational data would be obtained from trained observers and perhaps also from target students and incentive recipients.

Existing Records. Typical school records would provide such relevant quantifiable information about target students as attendance, homework and school work assignments completed, and percent of correct answers on daily quizzes. Since teachers differ in the amount and difficulty of their assignments and tests and in their criteria for grading students' work, it would be desirable to obtain baseline measures of these variables, utilizing records from past years.

Self-Reports. Information about the use of innovations and any formalized incentive system could be obtained by self-report procedures. Such procedures would be designed to elicit explicit information about teaching personnel, teaching methods, and teaching resources. Separate self-report measures would be designed for teachers, peer-tutors, and parents. The teachers' self-report measures would be the most elaborate.

Questions about teaching personnel would indicate the use of: contracted personnel with special expertise, specialists within the school system, team teaching, and teaching aides. The questions would cover special training, division of responsibilities and amount of time devoted. Especially important would be questions which indicate the quantity and type of teacher-to-teacher interactions outside of the classroom relative to the incentives experiment.

Questions about teaching methods would indicate the use of individualization of instruction and motivation techniques. Information about the use of incentive systems would be obtained. The latter might include the use of a token or point system, student contracts, reinforcement menus, setting of short and long range goals, modeling, and so forth. The relative amount of student versus teacher direction would be ascertained both in individual and in small group activities.

Questions about teaching resources used would indicate use of such media as programmed texts, film media, mechanical devices such as audio and video tape recorders, educational television, flash cards, workbooks, and so forth. An attempt would be made to determine not only their availability but also their use.

Although provision would be made for some open-ended responses, this would be minimized. Perhaps a small sample of interviewees would be questioned in depth to permit appropriate probes during the information acquisition process. Nevertheless, emphasis would be given to stimulus items that permit discrete, objectively determined, unequivocal responses. A few possible sample items of this variety are presented below:

1. Did you share the teaching of math with any other teacher?

   ___ yes ___ no If yes, please answer 1a through 1c.

   a. How many teachers were on the team? ____

   b. How were instructional responsibilities divided between (among) you (check all that apply)?

      ___ We planned over 50% of our lessons together.

      ___ We presented over 50% of our lessons together.

      ___ We divided the math topics between (among) us and each had primary responsibility for his topics.

      ___ We offered each other suggestions about how to improve instruction. If so, about how many times was this done?

         ___ less than 10 times

         ___ 10 to 20 times

         ___ more than 20 times

   c. Did this team receive special training in team teaching?

      ___ yes ___ no If yes, what kind (check all that apply)?

      ___ We each read books on the topic.

      ___ At least one of us attended a workshop on the topic.

      ___ A school supervisor provided consultation.

      ___ An expert from outside the school provided consultation.

      ___ Other: (please specify) _____

      _____

Observational Data. In contrast to self-report data, observational data are reported by one person about another. Such data may be obtained by trained outside observers. Also, participating teachers might report observations of the target students and, conversely, the latter might report observations of their teachers. Observational data may be recorded from memory; isolated incidents may be recorded when they occur;

or ongoing events may be recorded at the time and in the sequence in which they occur. The most useful information about cognitive behaviors, about the use of praise and punishment, and about the amount of student study behavior, would probably be obtained by trained outside observers recording classroom processes during the normal course of their occurrence. Such a data source would be relatively objective, it can be easily checked for observer reliability, and it would provide an estimate of the amount of time spent in each category of behavior.

There is no single observational scale that would provide all, but no more than, the information of interest to this study. An instrument for recording "cognitive behavior" (Webb, 1970) has been developed which was based upon the taxonomy of educational objectives in the cognitive domain (Bloom, 1956). The 55 categories of the Taxonomy of Cognitive Behaviors are subdivided into nine major categories that purport to range along a continuum from simple to complex. The nine categories are: knowledge of specifics, knowledge of ways and means of dealing with specifics, knowledge of universals and abstractions, translation, interpretation, application, analysis, synthesis (creativity), and evaluation. It may also be useful to obtain information from other dimensions of classroom behavior outlined in the taxonomy of teacher classroom behavior developed by Openshaw and Cyphert (1967).

One of the simplest procedures for recording a student's study and non-study behavior and teacher attention to student behavior is that reported by Hall, Lund, and Jackson (1968). Brophy and Good (1970) developed a more complex instrument to assess teachers' communications about their expectations for children's performance.

These and many other existing scales would provide the ingredients for an observation scale that would serve the purposes of this study. This hybrid scale would provide for the recording of both the incentive recipient and the target student behavior or perhaps the interactions of the two, on a time sampling basis.

GENERAL DESIGN CONSIDERATIONS

Included in this section are general design considerations which would apply to an incentives experiment regardless of which models were chosen to constitute the experiment from among those presented in the next chapters.

These models would all utilize a formal system of incentive delivery, in that required outputs would be identified in advance. Within this context, degree of prespecification of incentives (i.e., variation in the nature and delivery schedule of the actual incentive) could be varied experimentally. All proposed incentive models would be positive in nature, in that incentives would be considered as rewards for demonstrated student performance gains rather than as sanctions for the lack of them. It is well recognized, however, that public failure to "deliver the goods" on the part of a participant could function as a strong negative incentive. The implications of this fact could be taken into account in making inferences from the experimental data.

In all cases the welfare of the participants in the study should come first. Detected adverse effects on children should be cause for immediate revision of an experiment.

At all sites, cooperation and rapport should be established between the responsible parties at all levels, including parents, teachers, administrators, school board officials, experimenters, and USOE personnel. It is critical that a high degree of communication be maintained at all times. The experimenter's staff would be responsible for developing procedures, including printed materials and audio-visual presentations, to explain the purposes and methods of the experiment. In addition, an on-site consultant representative of the target population should be retained and trained to maintain good information flow between community, school personnel and the experimenters.

In attempting to conduct the most effective program that is possible without causing public concern or resistance, the experimenter should also provide local advisory groups that would be composed of persons in the community interested in improving educational outputs. The groups would review the program objectives and could serve as a third party in establishing measurable criteria for incentive delivery.

The proposed experiment is designed assuming the following degree of control over the experimental situation:

(1) A "subject pool" can be identified such that past biographical and educational data, including some form of psychometric data, exist and may be readily obtained for each potential target student.

(2) Classrooms or students can be randomly assigned to experimental (incentives manipulated) and control (incentives not identified) conditions.

(3) At least pre and post experiment data can be collected from participating experimental and control classrooms.

(4) A considerable amount of "process" information can be collected from experimental classrooms (see previous chapter devoted to this issue).

Since this is to be school-based rather than a laboratory expeirment, it is necessary to make some of the above assumptions without assurance that they can be met. Before embarking upon a larger scale experimental program, however, it is essential that a set of procedures for establishing these assumptions be devised, tried out, and evaluated. It is proposed that a small scale pilot study, containing selected elements of the proposed experiment, be developed and run in the spring of 1971. This would allow the refinement of subject identification and assignment, data collection, data analysis, public relations, and budget monitoring techniques prior to, rather than during, the projected experiment.[1]

It is proposed that in cases where the situation permits, the unit of experimental sampling should be the intact classroom, rather than the individual student. Many problems in the conduct of the experiment and in the analyses of data may be minimized through this strategy. It is felt that the consequent loss of statistical power is more than compensated by these advantages. However, some of the models may require that individual students serve as the experimental unit. This would be especially true of student and parent incentive models.

In all cases, some attempt should be made to measure for retention effects several months after the end of the incentive programs. It is possible that some aspects of the experiments could produce attenuated achievement

[1]A proposal for this pilot study has been submitted to USOE.

gains as measured by psychometric devices. For example, if the experiments involved large amounts of testing, improvement could possibly result simply from improved test-taking skills. Permanent gains should be sought, and verification of permanency by the experimental procedures is essential.

As was mentioned in the earlier discussion on criterion measurement, a large proportion of the output data will be collected in simple forms, such as the gain in percent of items answered correctly for a given objective over the course of the experiment. These data will be amenable to graphic presentation against a baseline of observations conducted both prior to the institution of incentives and after the incentive system has been removed. Within-site comparisons against the gains of well-defined (but perhaps not randomly selected) control groups will be made. Conventional non-parametric statistical analyses can be performed <u>if</u> randomization is achieved. These analyses could attribute statistical significance to obtained within-site differences.

The strongest case for demonstration of incentive effects will be made if incentive models at <u>different</u> sites with <u>different</u> subject populations produce similar positive results, as determined by <u>within-site</u> experimental versus control comparisons. It is well recognized that <u>between-site</u> differences may be expected to contribute a large amount of the variance in any across-site comparisons of the effectiveness of incentives models. For this reason, it is suggested that such comparisons be limited to examining degree of replication, based upon knowledge of important characteristics of each site. Such characteristics as race, socio-economic level, degree of site cooperation, grade level, geographical location, etc., could be indicators of site differences.

General Linear Hypothesis

Most research questions can be formulated as a specific case of the general linear hypothesis (see Bottenberg & Ward, 1963; Kelly, Beggs, & McNeil, 1969). This model is appropriate to the analysis of continuous psychometric achievement data in conjunction with dichotomous and discrete variables associated with membership in various groups. Such groups can be defined on the basis of data collected prior to or during an experiment. It can then be determined whether knowledge of these group memberships accounts for more than a chance amount of the total variation in the achievement scores of the experimental population. Analysis of variance

and analysis of covariance are just two of the statistical procedures that are facilitated by this model. Computer programs now exist (Dixon, 1969) to facilitate such analyses, and it is expected that, if random assignment of subjects is achieved, these programs could be applied to some of the data generated by the incentives experiment.

Another useful regression-based procedure would be to compare achievement prediction equations which are generated using the data from different sites. If the equations differ, it will be of interest to examine the regression coefficients to suggest the ways in which the sites differed in their effects.

Cost-Effectiveness Analysis

To the extent that a single set of educational objectives can be identified for the various models, cost/benefit analysis (Carpenter & Haggart, 1970) should be attempted to indicate the relative effectiveness of these models across sites. An effort could be made to relate the costs of various inputs as determined by the process evaluation to various outputs as considered in the chapter on criteria measurement.

STUDENT INCENTIVE MODELS

Two potential groups of subject populations are subsumed under this category. Included is the basic model in which a student would be compensated or rewarded to some extent contingent upon his performance. Another model would involve the compensation of other students, who may play a supportive or tutoring role, based upon the performance of the target students. The first model has received a considerable amount of attention from educational researchers, as evidenced by the previous literature review. The possibilities of each of these models to serve in an experimental program are now covered in turn.

Students as Recipients--Microincentives Model[2]

Of all possible models, this one probably has the most potential for improving educational practice. Since student learning is the variable of most interest, and since learning is defined in terms of an observable and permanent change in student behavior, any technique which has demonstrated effectiveness in bringing about such a change must be considered first in designing an educational improvement program. There can be no doubt about the validity of the mountainous body of data supporting the use of incentives to change a wide range of student behaviors. Of particular interest is a less numerous set of studies which show how extrinsic incentives[3] may be used as part of a systematic remedial program to improve academic (as opposed to personal or social) performance of disadvantaged and underachieving

[2]The term "microincentives" will be used henceforward in this report to differentiate incentives which are contingent upon small observable increments in student behavior, as opposed to "macroincentives," which will be used to refer to incentives which are contingent upon very large units of student performance, such as achievement test gains from September to May administrations.

[3]Extrinsic incentives, in this case, are thought of as identifiable consequences of responding which are delivered by mechanisms outside the cognitive system of the learner. These are as opposed to intrinsic incentives, which would occur as difficult-to-identify internal events within the learner, such as knowledge of success, enjoyment of reading, etc.

children. The most common academic behaviors studied are in the areas of reading readiness, reading, and mathematics. This research has generally involved a range of incentives, often utilizing a token delivery system, contingent upon small positive increments in performance. These performance increments are either observed by the teacher (or teacher aide) or are reflected in short mastery tests covering the content of brief instructional units. Such efforts, in some cases, have developed into well-defined educational programs characterized by sets of hierarchically organized student objectives stated in observable performance terms, training components for teachers, classroom organization procedures stressing individualization, etc. The importance of such programs for the present study cannot be overstressed.

In answer to the potential research question, "Are direct microincentives (regardless of type) to children effective in improving academic performance? " the available data seem to support a positive conclusion--given the amount of structure and reorganization found in such programs. Appendix A of this report presents up-to-date information on a number of educational programs which stress incentives to students. In light of this information, based not on small scale laboratory experiments but on large scale demonstrations involving thousands of public school children, a high proportion of whom represent low-achieving disadvantaged populations, the project staff have decided not to recommend the implementation of the model directly manipulating only microincentives to pupils. This is not to say that all possible knowledge exists about how to implement better incentive programs based upon small demonstrated performance gains. Many questions remain to be studied experimentally, including, for example, some of the following: Do physical incentives (such as candy, toys, reinforcing activities, etc.) as opposed to social incentives (such as adult approval) produce better results with low socioeconomic status youngsters? Are mechanical incentive delivery systems more efficient than teachers or paraprofessionals? Is group delivery of incentives more effective than individual delivery? Is immediate delivery of an incentive more effective or can small contingency contracts with children be utilized to elicit larger chunks of behavior before an incentive is granted? Can extrinsic incentives be "faded out" over time, allowing the intrinsic incentives contained in the instructional process to take over and

maintain a high frequency of responding on the part of the learner? It is felt that questions such as these will be answered within programs of research currently going on in numerous centers around the country.

To briefly recapitulate what has been said, it is felt that incentives granted directly to learners based upon small demonstrated performance gains are a very promising development in terms of improving educational practice. Incentives can enhance student performance on the rather tedious individual learning trials required to master skills such as reading and mathematics, thereby enabling the student to avoid becoming a functional dropout at an early age. Once he has mastered the skills requisite to <u>acquire</u> knowledge independently, some of the higher order incentives which exist <u>within</u> the educational process, such as opportunity to succeed and opportunity for self-direction, come within his grasp. Ultimately, this means the ability to function as a member of our economic-success oriented society. It is felt that educational programs currently exist, or are under development, to take advantage of this potential. While scientific questions still persist as to the most effective methods of microincentive utilization with student populations of various characteristics, the overall question of effectiveness has been answered in the affirmative. The expenditure of federal funds through the present project to further clarify this issue does not seem justified, especially in view of the substantial revisions in the "conventional" educational program required to implement such a system.

A possible alternative strategy would be to undertake a concentrated data collection and dissemination effort regarding the procedures and effects of the largest student-incentive based programs, paying particular attention to incentive aspects. Appendix A of this report is a first attempt at such a document, but it is realized that neither extensive selection procedures nor extensive site data collection efforts are represented therein.

Students as Recipients--Macroincentives Model

The distinction was previously made between microincentives, characterized as reinforcement for small performance gains, and macroincentives, referring to incentives which are made contingent upon very large units of student performance--such as would be the case where school-year achievement test gains are the criterion of interest. In direct contrast to the former type of incentive application, the latter approach has virtually no research history upon

which to draw. Several reasons seem apparent for this lack. Most important, there is room for serious doubt about the value of any technique which defers reinforcement for a period of months. Microincentives, in addition to serving a motivating function, act also to direct the instructional process by maintaining high rates of response on activities related to educational objectives. Macroincentive delivery, separated by a considerable lapse of time from the learning upon which it is contingent, would be unlikely by itself to either motivate students' performance or direct their efforts toward appropriate activities. This is especially true since the achievement measured by the typical standardized reading and math instrument is probably cumulative and takes an extremely large number of individual learning trials to acquire (see Staats, 1968).

Nevertheless, under certain conditions there is reason to expect that a student macroincentive model might have highly desirable results. These conditions can be described in more detail, given more knowledge about the characteristics of a particular school situation. It is assumed that they will have a higher probability of being effective with older students, but this may not necessarily be so. In general, the concept is that some material incentive, such as a trip, tickets to some desired event, or altruistic contribution to a needy community service, would be made contingent upon some well publicized gain in level of group performance. Of special interest would be problems of determining the most functional incentives. Techniques for fostering student group involvement would be utilized, and better learning habits, and more study, etc., would be encouraged by group pressure throughout the period during which incentives were offered. Teachers would be encouraged to cooperate by offering extra help and tutoring to those who desired it.[4] An aura of spirit typical of current interscholastic athletic competition could be sought.

In cases where sequential sets of behavioral objectives can be identified and tests for their attainment can be developed, it may be possible to identify shorter intervals and allow for more frequent incentive delivery. Within each experimental site it should be possible to try out several intervals and determine their relative cost-effectiveness.

[4]It might be advantageous to point out to the group that the most effective method to achieve a drastic increase in achievement levels would be to improve the performance of the very lowest group of students whose influence would have a disproportionate effect in lowering the group average.

Peer Tutoring

Peer tutoring is an educational technique just coming into limited use in public school settings. Generally, the technique allows students who have demonstrated mastery of a concept to teach other students who have not yet achieved such mastery. This requires a fairly sophisticated educational system in which diagnostic testing and individualization of instruction exist. Preliminary data from such systems indicate that within such constraints, the opportunity to serve as a tutor is itself an incentive for learners to tackle difficult learning tasks (see Niedermeyer and Ellis, 1970).

Considerable overlap exists here with systems of microincentives for students. It is difficult to conceive of appropriate conditions for delivery of incentives to a peer tutor without some companion system for measuring small performance gains. Moreover, given such a system, a gain itself is likely to act as a stronger incentive for the concerned tutor than just about anything else.

It is possible that peer tutoring could be profitably made a part of the macroincentive model that was presented above. Special recognition and help could be devised for those interested in participating as tutors. This merges somewhat with other techniques for securing group involvement. No identification of a peer model seems justified, in view of the desire of the present study to examine the effects of incentives without requiring large changes in the instructional program.

SCHOOL PERSONNEL INCENTIVE MODELS

It is generally accepted that student learning is the most important output of the educational process. School personnel, as a very important part of the process, presumably function with a primary goal of promoting learning. It would stand to reason that any technique which measureably improves student performance should thus serve as an incentive for all school personnel. Indeed, this is a tenet often cited by spokesmen of professional organizations such as NEA in arguing against merit pay systems (see Education Turnkey News, September, 1970). By a somewhat analagous chain of reasoning, proponents of reinforcement (microincentive) systems for children have generally deemed it unnecessary to devise special extrinsic incentive systems for teachers, teacher-aides, and administrators. The techniques of such systems literally force attention to student behavior--specifying, recording, and charting it (see Appendix A). Under such conditions, improvement or failure to improve is easy to recognize. It may well be, then, that student incentive models, if properly implemented, would make it unnecessary to look at further incentive models for school personnel. The authors of this report are not convinced that this is the case. The review of literature presented earlier points out several outstanding examples of drastically increased student performance which could be directly attributed to increases in external incentives to a responsible school figure. A line of research from another area has demonstrated possible mechanisms for this improvement. Rosenthal and Jacobson (1968) presented data to suggest that teacher expectations about student achievement had a subtle but demonstrable effect on actual achievement. Although Robert Thorndike (1968) presented a devastating rebuttal to this inference, citing methodological problems in the data-gathering operations of Rosenthal and Jacobson, recent research which appears methodologically sound (Brophy and Good, 1970) has produced similar results and has pointed out behavioral mechanisms involved in the function of teacher expectations as self-fullfilling prophecies. To quote these authors, "...the data suggest that the achievement levels of (these) classes were related to the teachers' performance demands and expectations." On this basis alone, there seems to be sufficient rationale for proposing the implementation of the following models.

It would be naive to assume, even in this period of increased militancy and concern for individual well-being, that school personnel are aloof from matters of their own personal effectiveness and the consequent intrinsic rewards. Often, however, some incentive must be offered in order to motivate an individual to expend the added effort necessary to break out of the status quo. As was postulated earlier in the case of children, once new and effective skills have been mastered, their use may provide enough intrinsic incentive to the teacher or administrator to maintain new behaviors in the absence of external incentives.

One cautionary note is especially relevant here. These models must deal with changes which will come about in the behavior of school personnel to influence and produce the changes in student performance, upon which incentive delivery will be based. When talking about student incentive models, there is a fairly high degree of assurance that behaviors which are rewarded will be the desired ones, or at least approximations thereof. When dealing with incentives to school personnel based upon student gains, there is no such assurance that desirable behaviors will be rewarded. Test coaching, unreasonable demands on children or teachers without proper foundations, communication and sharing breakdowns between teachers, etc. are just samples from the wide range of undesirable behaviors which could be rewarded through these techniques. It is important that safeguards against such outcomes be built into the system. If they are not, short term gains could fade and students over the long run could be negatively influenced. A separate chapter has previously been devoted to this issue.

Teacher Incentives Models

The most controversial potential application of an incentive system is also the least well researched, this being a situation in which teachers would receive direct extrinsic incentives based upon the performance of their pupils. Available information about the widely publicized failure of "merit pay" systems contains very little in the way of empirical findings about the effects and form of these programs. An examination of the characteristics of some currently operating teacher merit incentives systems (Weber & Marmion, 1969) reveals that most are monetary in nature and are based on ratings of "teacher quality" rather than any direct measures of student achievement. This suggests that difficulties in determining criteria for "merit" are

important contributors to the prevailing attitude of rejection. Nevertheless, there is every indication that public sentiment for performance incentives (cf. Gallup, 1970) will require some sort of cost accounting based upon educational output.

Other types of teacher incentives have been identified by designers of differentiated staffing programs (Weissman, 1969; Rand & English, 1968). These include such items as increased responsibility and recognition of achievement. As pointed out in the previous literature review, there are data to suggest that these factors, coupled with financial rewards, can modify ineffective teacher behavior and can serve to keep effective teachers in the classroom (as opposed to current procedures which tend to attract them to administration positions, where they may fail, cf. Peter & Hull, 1969). In general, three models are proposed here, differing mainly on the extent to which they are based on rewards for individual versus group participants.

Competitive Model

On one extreme is the model in which a school board or other responsible educational agency will agree to pay performance bonuses to individual teachers whose students achieve beyond some expected level. The expectations and contingencies of payment will be set out clearly at the beginning of the school year. Participation in the program should be at teacher option, meaning that teachers who do not wish to be considered for bonuses would not be required to apply, but applications would be necessary soon after the opening of school. Participating teachers should be allowed to draw advances against these bonuses to provide materials which they feel would be helpful in meeting the goals of the experiment, but no organized set of instructional procedures should be imposed as part of the experiment, and bonuses could be spent in any way the teacher desired. Responsible educational agencies could justify the expenditures for individual teacher incentives on the basis of increased cost effectiveness.

This model must be closely monitored to detect the nature of the educational changes undertaken by participating teachers, since the methods of producing increased results will ultimately be of more interest than the incentives which are used. Careful measurement should also be attempted of all outputs of the educational process, to detect whether rewarded outcomes are achieved at the expense of other desirable outcomes. Care should

be taken to detect adverse peer relationships which might result among teachers, but there is no a priori reason for assuming that bad feeling would result, since presumably all participating teachers could receive maximum bonuses.

A possible application of this first model especially suitable to a situation where extra money is not available for incentives is suggested by a program currently being tried out in San Mateo, California (see Weber & Marmion, 1969, p. 27-28). Under this application, teachers whose students make large gains beyond those that were expected could be elected with a certain amount of favorable publicity to a District Academy of Instruction. Although no additional stipend would accrue to teachers who qualified as members of the Academy, the prestige and recognition of membership could serve as powerful incentives. It might prove fruitful to have this body serve some advisory function for other teachers in the school, since the members presumably would be qualified through their demonstrated proficiency in bringing about gains in student achievement. In that case, additional funds could be provided to support these activities.

Cooperative Model

A second distinct model would compensate teachers as a group for performance gains on the part of their classes. This would be similar in most respects to two current projects underway as part of the OEO national performance contracting experiment. These "incentives only" projects in Mesa, Arizona, and Stockton, California, represent conditions in which the local professional teacher groups contract with the Boards of Education to produce specified achievement gains. Teachers are free to divide their bonus fees as they see fit, with current indications being that funds will be reinvested in incentives for students and in better educational materials for the schools. At present, not a great deal is known in Mesa or Stockton about the operational details and experimental sophistication of these projects. Efforts will be made to gain such information as decisions emerge. The AIR project staff feel that further experiments of this nature are well justified at other sites, examining important variables that may not be receiving attention in the OEO experiment. In particular, analyses should be made of the effects of non-monetary incentives in this type of situation; for example, opportunities for teachers to initiate their own programs through increased release time may well be more influential than the preoffered bonuses in bringing about productive changes.

Group Participation Model

The third model would attempt to go as far as possible in securing teacher group cooperation in the diagnois of student needs and prescription of remedial programs. A wide range of student needs, including but not limited to reading and math deficits, should be identified by teacher/community/ school administration action. Through similar techniques, funding requirements should be determined for implementing various strategies of meeting these needs and success criteria accepted by all parties. Teachers would then cooperate to bring to bear the best available materials and techniques for achieving these criteria, taking into account the fact that acceptance of responsibility and control is correlated with commitment to objectives. Demonstrated achievement would be required before incentives could be granted, but, again, advances could be allowed. Considerable teacher cohesiveness and preparation would also be required; this could not be guaranteed in most situations. However, in the opinion of the project stff, such a model has the greatest potential payoff in terms of making the best possible educational program compatible with the concept of accountability.

Teacher-aide (Paraprofessional) Model

It is felt that the current trend toward paraprofessional educational personnel suggests a potential population for incentive delivery. As in the case with peer tutors, however, it is recognized that paraprofessional aides exist mostly within the confines of educational systems which are already innovative to some extent. Nevertheless, it may prove possible to identify situations, in which paraprofessionals have some specific instructional responsibilities, where extrinsic incentives can be granted to some aides and withheld from a similar control group. In essence, it might be possible to create on a small scale the same experiment that would be going on with teachers under one of the first three teacher-incentives models. A similar demonstration of effect might be achieved with more control over the experimental situation and with smaller units of criterion measures. Very much depends upon the characteristics of the available population and site.

Administrators as Recipients

Considered here will be situations in which principals, superintendents, or entire school districts would be rewarded on the basis of student achievement gains above some expected level, based upon measured student ability and past performance.

Depending upon the pattern of staff arrangement prevailing in a school system, it is possible that principals could be either included with teachers in the incentive models presented previously or considered apart. On the basis of all that is currently known about the newly emerging role of building principal, it seems that the former arrangement would have most potential. Although he is certainly responsible for a large percent of the variance in such measures as school climate and staff morale, convincing data have not been found which can separate the principal from his staff in terms of influencing student achievement. Thus while it is difficult to justify a separate principal-incentive model, it seems wise to include principals in the teacher models.

Incentives to central school districts present an entirely different problem. In the first place, it is interesting to speculate about who would deliver the incentives at the district level, and upon what criteria delivery would be perpetrated. Assuming that a state or federal level agency could be influenced to play the role of monitor and dispenser, incentives for greater gains on achievement tests would serve simply to make the rich richer and the poor poorer. Two models proposed by Singell and Yordon (1970) based upon the Colorado Plan for Contract Accreditation seem to have possibilities, although the present authors would request clarification at several points and possibly suggest several modifications. In essence, these models propose that educational goals for students be set by districts through a process stressing community and teacher involvement. The state would then grant accreditation or additional funds based upon the success of a district over a three or four year period in meeting its goals.

There seem to be no assurances as to the overall excellence of these goals. Nor are criterion measurement techniques suggested beyond standardized achievement test gains. Under these conditions there is every evidence that middle socio-economic class districts could make the greatest gains and would hence be eligible for greater incentives than would lower socio-economic class districts. Singell and Yordon's suggestion that their model would "allow only school systems that have a demonstrated ability to improve students achievment to engage in innovation (1970, p. VII-5)" seems to be <u>opposite</u> the proper direction. Schools which are <u>not</u> successfully innovating are exactly the ones which need greater incentives based upon gains in student performance.

It is obvious that considerably more thought will have to go into the design of experiments for incentives at the district level.

PARENT INCENTIVE MODEL

Considerable interest has been generated recently by laws passed in several small communities in the Midwest rendering parents punishable for the crimes of their offspring. There can be no question that parents are largely responsible for creating the initial behavior patterns of their children. This is not the place for a discussion of the possible methods of this influence; however, it is safe to say that a large task of the school system is to help children acquire behaviors which may or may not be reinforced later in their home setting. Providing parent incentives along with instructions on the type of student behaviors to foster seems to have high potential for producing gains in student performance. Evidence from the area of social work with adolescents (Stuart, 1970b; 1970c) and preschool education for reading readiness (Niedermeyer, 1970) supports this hypothesis.

As has been pointed out earlier, there can be no question about the intrinsic incentives that accrue to most parents based upon demonstrated academic success of their children. What is needed, perhaps, is some external incentive for parents to try out techniques which have a higher probability of producing sucess. Since parents (as opposed to professional educators) cannot be assumed to have any knowledge of such techniques, some type of parent instruction will be a prerequisite for the implementation of an experiment in parent incentives. The more directly such incentives can be made contingent upon specific student behavioral changes, the greater the likelihood of success. Sets of parent training materials similar to those produced by the Southwest Regional Laboratory for Educational Research and Development (1970) in Los Angeles should serve this function, since they bear a clear relationship to academic improvement.

Teacher cooperation will be an important prerequisite to the implementation of a parent incentive program, since teachers would serve as the most logical contact for explaining the objectives and methods of an experiment to parents. The academic objectives of the program would be spelled out well in advance in a series of parent-teacher conferences and some type of group instruction on parent tutorial methods could be provided. Teachers who participate should also be allowed release time or other compensation in consideration for this assistance.

SUMMARY OF MODELS

Six different experimental models have been suggested in the previous three chapters with indications for a seventh. Two models involve students as incentive recipients, three definite models and a fourth possible model involve school personnel as recipients, and one model involves parents.

Within the area of student incentives, microincentive and macroincentive models were identified. It was recommended that only the latter be included in a large scale experimental investigation of incentives effects. Competitive, cooperative, and group participation models were identified in the area of school personnel incentives; a fourth model might involve incentives to paraprofessional aides if such personnel were found to exist within sites interested in participating in a large scale experiment. Finally, a parent incentive model was described. These six models[5] will serve as the components of the incentives experiment to be suggested in the following chapter of that title.

To summarize what has been said about the implementation of these models, there are abundant data to suggest that by continually reinforcing small student performance gains with a wide range of incentives, under conditions where instructional stimuli are well sequenced and measurable criteria of success are prespecified, considerable gains can be expected over conditions where such a combination of events does not occur. There are no data of which the authors are aware to suggest that simply paying students or school personnel or school districts based upon gains in composite grade equivalent scores will result in improved educational outputs. This is certainly an empirical question open to verification by experimentation; in general, however, it has been suggested that the best results of using extrinsic incentives are likely to be obtained when potential participants have some idea about what they need to <u>do</u> to improve the skills defined by sets of student performance objectives. Therefore, all proposed incentive models have included some provision for pointing out to or eliciting from the recipient populations techniques for better <u>achieving</u> specified educational outputs. This seems far more important to the authors, as educators, than the determination of how monetary incentives will be <u>used</u> by target population after they are earned.

[5] Student macroincentive; teacher competitive, cooperative, and group involvement; paraprofessional; and parent

Second, the provision of rewards contingent upon demonstrated student achievement gains assumes that something can be done differently from the way it is currently being done. It is very important to provide the climate and/or funding, as part of the experimental condition, through which changes can be allowed to occur. This means that, if cash incentives are utilized, professional educators should be allowed the opportunity to draw advances against whatever gains they feel they are capable of generating. If monetary incentives are not utilized, it is still crucial that the opportunity to change materials or techniques either through new training or personally-determined research utilization, be granted. It is not feasible to offer formal training programs or workshops to participating teachers as part of the incentive models, due to the "contamination" that such training would have on incentive effects; however, all participants should be allowed to have access to research such as that referenced in the literature review as part of orientation to the project.

Finally, the incentives suggested previously that can be utilized within the context of these models will involve various levels of monetary investment. In every case, however, the funds for providing incentives can be justified on the basis of the increased and demonstrated educational outputs. Singell and Yordon (1970, pp. VII-1 to VII-3) have presented an interesting economic analysis in this area based on increased tax revenues derived from students who stay in school longer. This point should be amplified at every opportunity to those that would decry such an investment. In reality, such an investment is the ultimate form of accountability.

AN INCENTIVES EXPERIMENT

Six experimental models have been suggested in which incentives would be offered to students, teachers/administrators, and parents based upon student attainment of well defined performance objectives in the areas of reading and mathematics. Each model represents the application of incentives in a situation from which very little in the way of empirical evidence has been generated to date. Each, however, in the judgment of the present authors, contains great potential for improving the practice of education in an area that represents one of the greatest national needs, i.e., improving the basic reading and computational skills of low achieving children. The basic approach of each model, moreover, utilizes the ultimate form of accountability, in that an effort would be made to define specific objectives and directly evaluate the degree of attainment of these objectives; no incentives could be earned unless demonstrated gains were achieved.

The problems of establishing appropriate educational objectives for students and functional incentives for diverse target populations are such that a priori theorizing and laboratory-based studies can only be suggestive of the direction for national policy in utilizing incentive techniques. As was pointed out in the introduction to this report, the potential for these models cannot be evaluated without an attempt to establish operational versions. Therefore, the authors have suggested that this attempt be made. The results of the first iteration must be considered as tentative until they have been revised in the light of experience. Schutz (1970) has pointed out the problematic nature of research on such complex educational processes. He has wisely advocated that it is more productive to view such research as "uncertainty reducing" rather than hypothesis testing. The present authors strongly feel that this should be kept in mind when reading the remainder of this chapter.

Possible Organization of a School-Based Experiment

Basic Experiment

The basic research question to be addressed by the proposed experiment is whether identifiable extrinsic incentives are effective in producing significant gains over and above control conditions where extrinsic incentives

are not preoffered. This question can be answered in as many as six different contexts utilizing the general design considerations (pp. 52-55) in conjunction with the individual models presented previously. For this reason, it is suggested that a variation of each model should be implemented in at least one location, yielding an experiment containing a minimum of six sites. The generalizability of such an experiment would be broadened if a heterogeneous and well defined sample of students at each site were selected to represent each target grade level (1-9). Approximately 1800 students would be selected at each site (900 receiving an incentive condition and 900 receiving "business as usual"), for a total estimated N of around 10,800 participants. Selection procedures would focus on the identification of a target population characterized by a considerable deficit in actual performance level when compared to their age/grade peers. The necessary degree of deficit would be set on an ad hoc basis to render the necessary number of subjects. Experimental (incentives manipulated) and control (incentives in the existing state) subjects, would both be selected from this group at random, if possible. The purposes of the experiment would best be served by identification of a single contractor to set up, monitor, and evaluate the models at all sites. This would probably allow better standardization of procedures than would be the case under multiple investigators. It would also allow the establishment of a central information bank on all subjects amenable to the needs of a coordinated data analysis. Appendix C contains tentative PERT networks which outline the suggestions of the AIR project staff as to the procedures and schedule for conducting the three basic models to students, school personnel, and parents.

As was noted in the chapter on general design considerations, across-site comparisons, necessary for between-model evaluations, would be quite difficult under this arrangement. Tentative conclusions based upon cost-effectiveness analyses in terms of attaining the same educational objectives might be possible; however, the judgment required in such analyses would be quite complex.

Combinations of Models

In addition, the basic question of incentive effectiveness could be

answered in any of 24 possible combinations[6] of multiple models. For example, student macroincentives could be combined with individual teacher incentives (competititve model) to produce potentially greater effects than would be generated by either model alone. Multiple models, however, make it difficult to isolate the effects of incentives attributable to the separate models. It is extremely unlikely that enough control could be exercised over a large scale field experiment of this nature to achieve the arrangement of subjects and models necessary to approximate a multifactor experimental design necessary to examine all potential main and interaction effects.

It is obviously of considerable importance, nevertheless, to estimate in some way the relative effectiveness of the models in isolation and in combination. One way to gather such information would be to attempt to implement in several sites an "additive" arrangement. Figure 1 illustrates this concept.

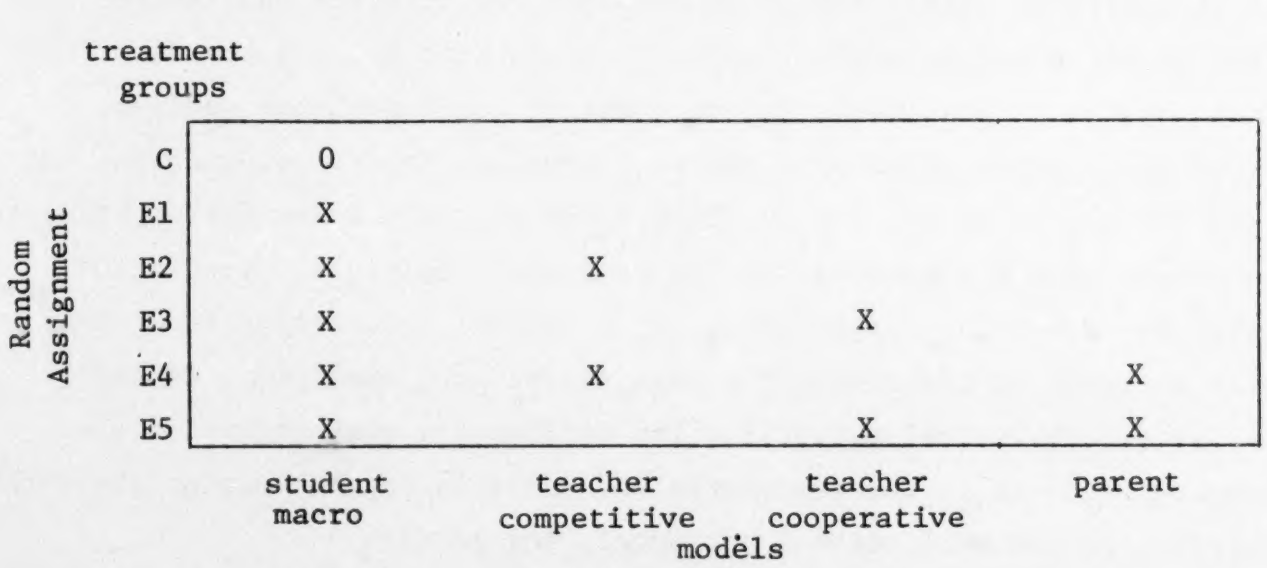


Figure 1.

Additive Design for Multiple Incentive Models

[6]Assuming three mutually exclusive teacher/administrator models, and one nonexclusive model each for students (macroincentives), paraprofessionals, and parents.

The particular design suggested in Figure 1 is a compelling one because of the emphasis on student incentives (as was pointed out earlier, it is felt that student incentives offer the greatest potential for enhancing student learning). Readers should note that such a design allows an estimation of the differential effects of teacher and parent incentive models when added to student incentives; no estimation of their effects in isolation, however, is justified. The logistical and administration costs of such designs would seem to preclude their establishment in all but the most favorable of sites. Considerably more subjects and funds would be required to implement a combination of models such as this.

Possibly the most attractive combinations, in addition to those illustrated in Figure 1, would involve: the student macroincentives model in conjunction with parent incentives alone and in conjunction with one of the four possible school personnel incentive models; parent incentives alone and in conjunction with school personnel models; the three basic teacher incentive models in opposition; etc. In all cases, the experimental versus control comparisons for each treatment combination would address the basic research question regarding incentive effectiveness.

Experimental Options

The United States Office of Education, as the recipient of this report, will have several options for supporting a school-based field experiment. The first is a go/no-go decision regarding the actual conduct of the experiment. If the decision is in the affirmative, much will depend on the funding options which are selected and on the receptivity of school systems around the country. The incentives for school systems, in addition to the potential improvements that might result in their educational program, are the external funds which will be made available to provide incentives. The next section is devoted to funding possibilities. Five school districts, sampled on a non-random basis as part of AIR's effort to determine the feasibility of a school-based experiment, submitted letters of interest which are contained in Appendix D. It is felt that, on the basis of this limited sample, enough schools can be found to permit conduct of the basic experiment.

School districts interested in participating in a field experiment should have the option of selecting one of the six basic models presented previously. They should then have the opportunity to select other basic models

or possible combinations of models, depending on local conditions and needs. Logistical problems in adequately evaluating an experiment of this type would dictate that probably not more than six sites be chosen for the field experiment. This decision will be reflected in all further discussions of the funding requirements, administration, etc. of the study.

The best possible situation would exist if six districts could be found which, in addition to possessing the requisite low achieving student populations and commitment to the study, would each be interested in implementing one of the six models. They could then opt for whatever other model or combination of the models seemed to meet their needs and potential capability to implement.

The matter of what type of incentives and what delivery schedule to use could also be made a matter of LEA option. The previous review of literature has covered both topics exhaustively. It is suggested that this issue be determined in negotiations with local officials and representatives of the potential target populations (in the case of teacher and parent models) prior to the experiment. Findings from a pilot study would be of great benefit at this stage for the large scale field experiment. In general, non-monetary incentives could involve such items as special recognition, opportunity to engage in valued activities, release time, etc. Monetary incentives, in addition to money, would involve the purchase of material items directly or through tokens such as trading stamps. In all cases, incentives could either accrue to the personal gain of the recipient or could be reinvested in the educational enterprise. Appendix E contains a 4 page document that might be suitable, after addition of appropriate funding information as an introduction to school districts. It explains on a preliminary basis the options that have been suggested above.

Funding Options

There are three apparent options for funding an experiment such as that proposed above.

The most frequently discussed option would involve funding from three sources--local, state, and federal. The cost of the instructional component would be borne by the participating school district. This would include teacher salaries, instructional materials, supplies, etc. The cost of

providing incentives would be supplied either by the local school district or to the district through state-controlled federal funds. The most likely sources of these latter funds would be Titles I, III, and possibly VIII of the Elementary and Secondary Education Act of 1965 (PL89-10 as amended); if large cities with considerable urban populations wish to become involved, it is possible that education funds could be allocated through the Model Cities program, under Title I of the Demonstration Cities and Metropolitan Development Act of 1966 (PL89-754 as amended). It must be kept in mind that even "non-monetary" incentives such as release time, trips, etc. would cost a certain amount of money. Such funds must be sought by local school districts through a proposal directed to the state educational authority or urban Model Cities administrative council. Finally the cost of the set-up, monitoring, and evaluation of the results of the experiment would be provided by evaluation funds controlled by the Office of the Assistant Secretary for Education or the Office of Program Planning and Evaluation. Such a funding consortium would be awkward at best, requiring considerable sophistication in securing necessary authorizations for expenditures at the various levels. It seems likely that coordination of incentive conditions would be quite difficult under circumstances where different authorities control these funds.

A second option would contain essentially the same sources as the first except that costs of set-up, monitoring, and evaluation would be included in proposals submitted to state education authorities under the various ESEA Titles. This would mean in all probability a series of self-contained experiments with less potential for overall coordination by a single agency. The findings of such experiments would nevertheless be of considerable interest and fund authorization problems would be minimized.

The third option would involve direct federal involvement in providing both incentives and evaluation. The incentives costs could be provided through ESEA Title III funds which have been assigned to the discretion of the Commissioner of Education. Management support, monitoring, and evaluation could be provided through evaluation monies available to USOE. This option would provide allocative <u>and</u> experimental coordination. It can be imagined that local school districts would be more cooperative under this option, in view of the fact that they would not be required to construct and submit proposals to the state education agency.

It is obvious that the opportunity to select funding options rests first with the U. S. Office of Education and second with interested local school districts.

Publicity

As a final note, it is apparent that one of the biggest problems associated with current incentive experiments, such as the OEO "performance contracting experiment," is too much publicity too soon. Verbalized hope coupled with a bit of journalistic flair can sometimes produce overinflated claims and misinformation which doom projects to "failure" from the beginning. Even though there is great interest in incentives projects such as those which are proposed in this report, the authors would urge that, should such projects become operational, publicity be directed mainly through professional channels until the results are in. There can be nothing more frustrating to the professional scientist than trying to dampen the enthusiasm of zealots jumping on an innovation bandwagon prior to the analysis of the evidence.

PROJECT COST ESTIMATES AND CONTINUATION OPTIONS

Under the options described in the previous chapter it is possible to conceive of several possible cost breakdowns for the field experiment. Assuming a six site study and a division of funding responsibility roughly equivalent to option three (instructional program-LEA; administration and incentives-Title III discretionary; management support,monitoring, and evaluation-USOE), the following estimates can be offered.

Table I presents the estimates for providing mangement support, monitoring, and evaluation through a single agency that would have responsibility for maintaining the experimental aspects of the program through the first year.

TABLE I

MANAGEMENT SUPPORT, MONITORING, AND EVALUATION

Set Up	
Salaries & benefits, including overhead	38,000
Travel	4,800
Communications	500
Fee	3,000
Contingency	1,000
Subtotal	47,300
Conduct Support and Monitoring	
Salaries & benefits, including overhead	114,048
Travel	14,000
Communications	1,250
Supplies	20,000
Services	4,000
Fee	15,000
Contingency	3,750
Subtotal	172,048
Analysis and Reporting	
Salaries & benefits, including overhead	38,000
Travel	1,400
Communications	250
Supplies	3,000
Services	6,050
Final Report Production	1,800
Fee	1,500
Contingency	500
Subtotal	52,500

Totals

Phase I	47,300
Phase II	172,048
Phase III	52,500
	271,848

Depending on which options are selected in the area of incentives, estimates for LEA-requested funding from Title III discretionary range from $100,000 to $150,000 for each site. The range is dependent upon the number of models a school district chooses to implement, the degree of monetary incentives selected versus special recognition incentives, the delivery method chosen, and the degree to which local support may be secured.

Local support would especially involve such things as facilities, clerical and logistical backup, cost differential between the salaries of professionals involved and the salaries of substitutes employed to temporarily handle their duties while they are participating in the incentives project, etc. In addition, it is possible that local support could be extended to the provision of incentives, either in the form of monetary contributions or by providing desirable activities or materials, contingent upon student performance gains. While it is possible to envision all manner of local contributions once an experiment is underway, it is difficult to specify such sources at this time, without knowledge accessible only to personnel within the target communities. If an incentives experiment becomes a reality, special efforts should be undertaken immediately in each participating site to locate and tap local incentive sources.

Continuation Options

It seems appropriate here to reflect back on two points which have been made previously in this report. The first is that the first year implementation of incentive models should be considered as an initial attempt, to be revised and improved on the basis of that experience. The second is that it would be desirable to study incentive effects on a longitudinal basis over a period of several years. This suggests the basic direction which all continuation options should take. A range of options is now presented.

Fallback option. At least one option should be available for situations in which the preliminary data on a field experiment are not encouraging. This situation could be envisioned if public relations problems, teachers union problems, communication breakdowns, etc. hindered the implementation of one or more models at a site. Also included would be situations in which incentives just were not producing the hoped for (and required) student performance gains on objectives in reading and mathematics. In such cases, it would be advisable to curtail expansion to any other sites and work with existing sites to remediate their problems for the second year's program. If problems seemed insurmountable, it might be necessary to eliminate sites and concentrate resources on other sites in which continuation seemed profitable, taking care to learn as much as possible about the causes of termination.

Modified experimental expansion. If general acceptance and results of the first year's field trials seem favorable, it is recommended that the trials be expanded on an experimental (as opposed to a fully operational) basis within the existing sites. This expansion would involve making modifications to the models as indicated by the on-going process evaluation, perhaps adding other models in other schools, adding more subjects within already implemented models, etc. The experimental aspect would be carried on throughout the second year of the program, at which time external incentives sources would be withdrawn and a third year evaluation would be undertaken on a longitudinal basis to determine the long range consequences of the incentive experiment to the educational programs of the original sites. Such an expansion would seem to offer the maximum payoff in terms of reliable information about incentive effects.

Operational expansion. This would involve a general dissemination of incentive techniques utilized in the first year of the field experiment. Ideally, such expansion would be delayed until after the third year of modified experimental expansion, but it could occur sooner if preliminary results were particularly encouraging. There would be little if any central experimental control over such expansion. Best possible data on outputs and techniques of the incentives experiment would be published and made available to interested schools. Some type of general management and technical support or handbook could be produced to help with locally controlled implementation and evaluation efforts. There would be no real limit to the range of expansion under this option. The current project staff are unanimous in the hope that this stage could eventually be reached.

BIBLIOGRAPHY

Allen, D. W. & Ryan, K. A. Microteaching. Reading, Mass: Addison Wesley, 1969.

Alschuler, A. The effects of classroom structure on achievement motivation and academic performance. Educational Technology, 1969, 9, 19-24.

Anderson, R. C. Educational psychology. Annual Review of Psychology, 1967, 18, 129-158.

Anderson, R. C. Control of student mediating processes during verbal learning and instruction. Review of Educational Research, 1970, 40, 349-369.

Ayllon, T. & Azrin, N. H. Reinforcer sampling: A technique for increasing the behavior of mental patients. Journal of Applied Behavior Analysis, 1968, 1, 13-20.

Badel, A. W. & Larsen, E. P. On reporting test results to community groups. NCME Measurement in Education, 1970, 1 (4).

Bandura, A. & Perloff, B. Relative efficacy of self-monitored and externally imposed reinforcement systems. Journal of Personality and Social Psychology, 1967, 7, 111-116.

Bandura, A. Principles of behavior modification. New York: Holt, Rinehart & Winston, Inc., 1969.

Becker, T. D. & Armstrong, M. Production and elimination of disruptive classroom behavior by systematically varying teacher's behavior. Journal of Applied Behavior Analysis, 1968, 1, 35-45.

Benowitz, M. L. & Busse, T. V. Material incentives and the learning of spelling words in a typical school situation. Journal of Educational Psychology, 1970, 61, 24-26.

Berlyne, D. E. Conflict, arousal and curiosity. New York: McGraw-Hill, 1960.

Bloom, B. S. (Ed.) Taxonomy of educational objectives: Cognitive domain. New York: David McKay Company, Inc., 1956.

Bottenberg, R. A. & Ward, J. H. Applied multiple linear regression. Technical Document Report PRL-TDR-63-6, March, 1963, 6570th Personnel Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Lackland Air Force Base, Texas. Project 7719, Task 771901.

Brophy, J. E. & Good, T. L. Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. Journal of Educational Psychology, 1970, 61, 365-374.

Bushell, D. Jr., Wrobel, P. D., & Michaelis, M. L. Applying "group" contingencies to the classroom study behaviors of preschool children. Journal of Applied Behavior Analysis, 1968, 1, 55-61.

Bushell, D. Jr. The behavior analysis classroom. Follow Through Project, Department of Human Development, University of Kansas, Lawrence, Kansas, 1970.

Campbell, V. N. & Chapman, M. A. Learner control vs program control of instruction. Psychology in the Schools, 1967, 4, 121-130.

Cantrell, R. P., Cantrell, M. L., Huddleston, C. M., & Wooldridge, R. L. Contingency contracting with school problems. Journal of Applied Behavior Analysis, 1969, 2, 215-220.

Carpenter, M. B. & Haggart, S. A. Cost-effectiveness analysis for educational planning. Educational Technology, October, 1970, 26-30.

Carver, R. P. Special problems in measuring change with psychometric devices. In Evaluative Research. Pittsburgh, Penn: American Institutes for Research, 1970.

Chadwick, B. A. & Day, R.C. Systematic reinforcement: Academic performance of Mexican-American and Black students. Unpublished manuscript, Washington University, 1970.

Clark, C. A. & Walberg, H. J. The use of secondary reinforcement in teaching inner-city school children. Journal of Special Education, 1969, 3, 177-185.

Clark, K. B. A design for the attainment of high academic achievement for the students of the public elementary and junior high schools of Washington, D. C. New York: Metropolitan Applied Research Center, 1970.

Cobb, J. A. The relationship of discrete classroom behaviors to fourth-grade academic achievement. Eugene, Oregon: Oregon Research Institute Bulletin, 10 (10), 1970.

Cohen, H. L. MODEL: Motivationally oriented designs for an ecology of learning. In H. L. Cohen, I. Goldiamond, J. Filipczak, & R. Pooley (Eds.) Training professionals in procedures for the establishment of educational environments. Silver Spring, Maryland: Educational Facility Press-IBR, 1968.

Coleman, J. S. Incentives in American education. Paper presented for Brookings Seminar on Incentives in Public Policy, February, 1969, John Hopkins University, Report No. 40.

Cormier, W. H. Effects of approving teaching behaviors on classroom behaviors of disadvantaged adolescents. Department of Educational Psychology and Guidance, University of Tennessee, Project No. 9-D-017, Grant No. OEG-4-9-520017-0029-057 with USOE, Bureau of Research, March, 1970.

Cotler, S. The effects of positive and negative reinforcement and test anxiety on the reading performance of male elementary school children. Genetic Psychology Monographs, 1969, 80, 29-50.

Cox, R. C. & Sterrett, B. G. A model for increasing the meaning of standardized test scores. Journal of Educational Measurement, 1970, 7, 227-228.

Cronbach, L .J. & Furby, L. How we should measure "change" - or should we? Psychological Bulletin, 1970, 74, 68-80.

Cronbach, L. J. & Snow, R. C. Individual differences in learning ability as a function of instructional variables. Final Report, March, 1969, Stanford University, Contract OEC 4-6-061269-1217, U. S. Office of Education.

Day, R. C. & Chadwick, B. A. Modification of disruptive classroom behavior of Mexican-American and Black children. Unpublished manuscript, Washington University, 1970.

Davidoff, S. H. The development of an instrument designed to secure student assessment of teaching behaviors that correlate with objective measures of student achievement. ERIC Document ED 039170, March, 1970.

Dixon, W. J. Biomedical computer programs. Los Angeles: University of California Press, 1969.

Education Turnkey Systems. Performance contracting in education. The guaranteed student performance approach to public school system reform. Champaign, Ill: Research Press, 1970.

Education U. S. A. Weekly newsletter on Education Affairs. Washington, D.C.: National School Public Relations Association, September 7, 1970.

Ferster, C. B. & Skinner, B. F. Schedules of reinforcement. New York: Appleton-Century-Crofts, 1967.

Flanagan, J. C., Shanner, W. M., & Mager, R. F. Behavioral objectives: A guide for individualizing learning. New York: Westinghouse Learning Corporation, 1971.

Flanders, J. P. A review of research on imitative behavior. Psychological Bulletin, 1968, 69, 316-337.

Fourth Annual Report. The National Advisory Council on the Education of Disadvantaged Children. Washington, D.C., 1969.

Fox, L. Effecting the use of efficient study habits. Journal of Mathetics, 1962, 1, 75-86.

Gallup, G. The public's attitude toward the public schools. Phi Delta Kappan, 1970, 52, 97-112.

Glaser, R. Instructional technology and the measurement of learning outcomes: Some questions. American Psychologist, 1963, 18, 519-521.

Green, R. L. & Stachnik, T. J. Money, motivation, and academic achievment. Phi Delta Kappan, December, 1968, 228-230.

Halcomb, C. G. & Blackwell, P. Motivation and the human monitor: I. The effect of contingent credit. Perceptual and Motor Skills, 1969, 28, 623-629.

Hall, R. V., Axelrod, A., Foundopoulos, M., Shellman, J., Campbell, R. A., & Cranston, S. S. The effective use of punishment to modify behavior in the classroom. Educational Technology, in press.

Hall, R. C., Lund, D., & Jackson, D. Effects of teacher attention on study behavior. Journal of Applied Behavior Analysis, 1968, 1, 1-12.

Harris, C. W. (Ed.) Problems in measuring change. Madison: University of Wisconsin Press, 1963.

Hart, B. M. & Risley, T. R. Establishing use of descriptive adjectives in the spontaneous speech of disadvantaged preschool children. Journal of Applied Behavior Analysis, 1968, 1, 109-120.

Hawkins, R. P. & Sluyter, D. J. Modification of achievement by a simple technique involving parents and teacher. Paper presented at the meeting of the American Educational Research Association, Minneapolis, March, 1970.

Hawkridge, D. G. Designs for evaluative studies. In Evaluative Research. Pittsburgh, Penn: American Institutes for Research, 1970.

Hawkridge, D. G., Chalupsky, A. B., & Roberts, A. O. H. A study of selected exemplary programs for the education of disadvantaged children. Final Report, September, 1968, Contract No. OEC-0-8-089013-3515 (010), Project No. 089013. Palo Alto: American Institutes for Research.

Hawkridge, D. G., Campeau, P. L., DeWitt, K. M., & Trickett, P. K. A study of further selected exemplary programs for the education of disadvantaged children. Final Report, June, 1969, Contract No. OEC-0-9-089013 (010), Project No. 089013. Palo Alto: American Institutes for Research.

Heitzman, A. J. Effects of a token reinforcement system on the reading and arithmetic skill learnings of migrant primary school pupils. Journal of Educational Research, 1970, 10, 455-458.

Helson, H. Adaptation level theory. New York: Harper Row, 1964.

Herman, H. & Adkins, D. C. Final Report: Hawaii head start evaluation - 1968-1969. Education Research and Development Center, College of Education, University of Hawaii, Honolulu, Hawaii, January, 1970.

Hewett, F. M., Taylor, F. D., & Artuso, A. A. The Santa Monica project: Evaluation of an engineered classroom design with emotionally disturbed children. Exceptional Children, 1969, 35, 523-532.

Homme, L. Contingency Management. Newsletter of the Division of Clinical Psychology, APA, 1966, 5 (4).

Jencks, C. Giving parents money for schooling: Education vouchers. Phi Delta Kappan, September, 1970, 49-52.

Kelly, F. J., Beggs, D. L., & McNeil, K. A. Research design in the behavioral sciences: Multiple regression approach. Carbondale: Southern Illinois University Press, 1969.

Kennedy, B. J. Motivational effects of individual conferences and goal setting on performance and attitudes in arithmetic. ERIC Document ED032113, July, 1968.

Kennedy, W. A. & Willcutt, H. C. Praise and blame as incentives. Psychological Bulletin, 1964, 62, 323-332.

Klein, S. Evaluating tests in terms of the information they provide. In Evaluation Comment, Center for the Study of Evaluation, University of California, Los Angeles, June, 1970, 1-6.

Krumboltz, J. D. & Wiesman, R. G. The effect of intermittent confirmation in programmed instruction. Journal of Educational Psychology, 1962, 53, 250-253.

Leith, G. O. M. & Davis, T. N. The influence of social reinforcement on achievement. Educational Research, 1969, 11, 132-137.

Leno, R. S. Children's methods of problem solving in arithmetic. (Dissertation submitted to the School of Education, Stanford University) 1958.

Lessinger, L. M. Every kid a winner: Accountability through educational engineering. New York: Simon & Schuster, 1970.

Lindvall, C. M. & Cox, R. C. Evaluation as a tool in curriculum development: The IPI evaluation program. Chicago: Rand McNally & Co., 1970.

Lipe, D. & Steen, M. T. Assessment of PLAN teacher development. Westinghouse Learning Corporation, Palo Alto, California, 1970.

Lovaas, O. I. Learning theory approach to the treatment of childhood schizophrenia. Behavior Theory and Therapy, California Mental Health Research Symposium No. 2, State of California, Department of Mental Hygiene, Bureau of Research, 1968.

Lovitt, T. C. & Curtiss, K. A. Academic response rate as a function of teacher and self-imposed contingencies. Journal of Applied Behavior Analysis, 1969, 2, 49-53.

Madsen, C. H. Jr., Becker, W. C., & Thomas, D. R. Rules, praise and ignoring: Elements of elementary classroom control. Journal of Applied Behavior Analysis, 1968, 1, 139-150.

Marshall, H. H. Learning as a function of task interest, reinforcement, and social class variables. Journal of Educational Psychology, 1969, 60, 133-137.

McAllister, L. W., Stachowaik, J. G.,Baer, D. M.,& Condermon, L. The application of operant conditioning techniques in a secondary school classroom. Journal of Applied Behavior Analysis, 1969, 2, 277-285.

McArthur, L. A. & Zigler, E. Level of satiation on social reinforcers and valence of the reinforcing agent as determiners of social reinforcer effectivenss. Developmental Psychology, 1969, 1, 739-746.

McClelland, D. C. Personality. New York: Holt, Rinehart & Winston, Inc., 1951.

McIntire, R., Davis, G., & Pumroy, D. Improved classroom performance by reinforcement outside of the classroom. Proceedings, 78th Annual Convention of American Psychological Association, 1970, 747-748.

McMains, M. J. Children's adoption of self-rewarding patterns: Verbalizations and modeling. Perceptual and Motor Skills, 1969, 28, 515-518.

Meacham, M. L. & Wiesen, A. E. Changing classroom behavior: A manual for precision teaching. Scranton, Penn : International Textbook Co., 1969.

Messick, S. Statement before the Select Subcommittee on Education of the Committee on Education and Labor, House of Representatives, Ninety-First Congress, 1969.

More, A. J. Delay of feedback and the acquisition and retention of verbal materials in the classroom. Journal of Educational Psychology, 1969, 60, 338-342.

Niedermeyer, F. C. Parent-assisted learning. Inglewood, California: Southwest Regional Laboratory for Educational Research and Development, 1970.

Niedermeyer, F. C. & Ellis, P. A. The development of a tutorial program for kindergarten reading instruction. Inglewood, California: Southwest Regional Laboratory for Educational Research & Development, 1970.

O'Leary, K. D., Becker, W. C., Evans, M. B., & Saudargas, R. A. A token reinforcement program in a public school: A replication and systematic analysis. Journal of Applied Behavior Analysis, 1969, 2, 3-13.

Openshaw, M. K. & Cyphert, F. R. The development of a taxonomy for the classification of teacher classroom behavior. Department of Health, Education, and Welfare, Cooperative Research Program of the U. S. Office of Education, Report No. 2288, 1967.

Osborne, J. G. Free-time as a reinforcer in the management of classroom behavior. Journal of Applied Behavior Analysis, 1969, 2, 113-118.

Packard, R. G. The control of "classroom attention:" A group contingency for complex behavior. Journal of Applied Behavior Analysis, 1970, 3, 13-28.

Panyon, M., Boozer,H., & Morris, N. Feedback to attendants for applying operant techniques. Journal of Applied Behavior Analysis, 1970, 3, 1-4.

Peter, L. J. & Hull, J. The Peter Principle. New York: Bantam Books, 1969.

Peterson, R. W. Accountability in elementary and secondary education. National Committee for Support of the Public Schools News. Washington, D. C., July, 1970.

Petrafeso, P. R. A specific educational role for the classroom aide. In Behaviorally Engineered Environments. Upper Midwest Regional Laboratory, Minneapolis, Minn., 1970.

Popham, W. J. The instructional objectives exchange: Progress and prospects. Center Reprint Series, No. 15. Center for the Study of Evaluation, University of California, Los Angeles, 1970.

Premack, D. Reinforcement theory. In D. Levine (Ed.) Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 1965.

Prince, A. J. Conditioning children for school. ERIC Document ED019802, December, 1967.

Rand, M. J. & English, F. Towards a differentiated teaching staff. Phi Delta Kappan, 1968, 49, 264-268.

Reynolds, M. J. & Risley, T.R. The role of social and material reinforcers in increasing talking of a disadvantaged preschool child. Journal of Applied Behavior Analysis, 1968, 1, 253-262.

Risley, T. R. The effects and side effects of punishing the autistic behaviors of a deviant child. Journal of Applied Behavior Analysis, 1968a, 1, 21-34.

Risley, T. R. Jenny Lee; Learning and lollipops. Psychology Today, 1968b, 1 (8), 28-31, 62-65.

Risley, T. R. & Hart, B. Developing correspondence between the non-verbal and verbal behavior of preschool children. Journal of Applied Behavior Analysis, 1968, 1, 267-281.

Rosenthal, R. & Jacobson, L. Pygmalion in the classroom: Teacher expectation and pupils' intellectual development. New York: Holt, Rinehart & Winston, 1968.

Ryback, D. & Staats, A. W. Parents as behavior therapy-technicians in treating reading deficits (dyslexia). Journal of Behavior Therapy and Experimental Psychiatry, 1970, 1, 109-119.

Schmidt, G. W. & Ulrich, R. E. Effects of group contingent events upon classroom noise. Journal of Applied Behavior Analysis, 1969, 2, 171-179.

Schutz, R. E. Programmatic instructional development. Inglewood: Southwest Regional Laboratory for Educational Research and Development, September, 1970.

Schutz, R. E. & Baker, R. L. "The experimental analysis of behavior" in educational research. Psychology in the Schools, 1968, 5, 240-256.

Schwenn, E. A., Klausmeier, H. J., & Sorenson, J. S. Motivational procedures in school settings: The effect of individual adult-child conferences in reading. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, Minnesota, March, 1970.

Schwitzgebel, R. L. Behavior instrumentation and social technology. American Psychologist, 1970, 25, 491-499.

Shanner, W. M. Evaluation of mastery level in systems of individualized instruction. In Systems in Education: A Seminar on the Usefulness of Systems Analysis in the Field of Education. Sponsored and published by Daniel, Mann, Johnson & Mendenhall, 1968.

Sidman, M. Tactics of scientific research. New York: Basic Books, Inc., 1960.

Simon, A. & Boyer, E. G. (Eds.) Mirrors for behavior. Philadelphia: Research for Better Schools, Inc., 1967, 1.

Singell, L. D. & Yordon, W. J. To develop alterative incentive models for education based on (but not restricted to) the Colorado experiment in accreditation of school districts by contract. Final Report, September, 1970, New York University, Contract No. HEW-05-69-94.

Skinner, B. F. The design of cultures. In R. Ulrich, T. Stachnik, & J. Mabry (Eds.), Control of human behavior. Glenview, Ill: Scott, Foresman & Co., 1966a.

Skinner, B. F. Operant behavior. In W. K. Honig (Ed.), Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, 1966b.

Smith, D. E. P. Maintaining task behavior in a literacy program under various conditions of reinforcement. ERIC Document ED015113, December, 1966.

Soloman, R. L. & Brush, E. S. Experimentally derived conceptions of anxiety and aversion. In M. R. Jones (Ed.) Nebraska Symposium on Motivation. Lincoln: University of Nebraska Press, 1956.

Southwest Regional Laboratory for Educational Research and Development. Parent-assisted learning program. Inglewood, California: SWRL, 1970.

Spence, J. T. A study of certain factors affecting children's school performance. ERIC Document ED011086, 1966.

Staats, A. W. Learning theory and "opposite speech." Journal of Abnormal and Social Psychology, 1957, 55, 268-269.

Staats, A. W. Learning, language, and cognition. New York: Holt, Rinehart & Winston, Inc., 1968.

Staats, A. W., Finley, J. R., Minke, K. A. & Wolf, M. Reinforcement variables in the control of unit reading responses. Journal of the Experimental Analysis of Behavior, 1964, 1, 139-149.

Staats, A. W., Minke, K. A., Goodwin, W., & Landeen, J. Cognitive behavior modification: 'Motivated learning' reading treatment with subprofessional therapy technicians. Behaviour Research and Therapy, 1967, 5, 283-299.

Staats, A. W., Staats, C. K., Schutz, R. E., & Wolf, M. The conditioning of textual responses using "extrinsic" reinforcers. Journal of the Experimental Analysis of Behavior, 1962, 5, 33-40.

Stuart, R. B. Behavior therapy techniques for use in the treatment of delinquents. Paper presented at the meeting of the American Psychological Association, Miami, September, 1970a.

Stuart, R. B. Behavioral contracting within the families of delinquents. Paper presented at the meeting of the American Psychological Association, Miami, September, 1970b.

Stuart, R. B. Trick or treatment: How and when psychotherapy fails. Champaign, Ill : Research Press, 1970c.

Sullivan, H. J., Baker, R. L., & Schutz, R. E. Effect of intrinsic and extrinsic reinforcement contingencies on learner performance. Journal of Educational Psychology, 1967, 58, 165-169.

Surratt, P. R., Ulrich, R. E., & Hawkins, R. P. An elementary student as a behavioral engineer. Journal of Applied Behavior Analysis, 1969, 2, 85-92.

Thomas, D. R., Becker, W. C., & Armstrong, M. Production and elimination of disruptive classroom behavior by systematically varying teachers' behavior. Journal of Applied Behavior Analysis, 1968, 1, 35-45.

Thorndike, R. L. Review of R. Rosenthal and L. Jacobson, "Pygmalion in the classroom." American Educational Research Journal, 1968, 5, 708-711.

Tyler, R. W. The objectives and plans for a national assessment of educational progress. Journal of Educational Measurement, 1966, 3, 1-4.

Tyler, R. W. The purposes of assessment. In W. H. Beatty (Ed.), Improving educational assessment and an inventory of measures of affective behavior. Washington, D.C.: Association for Supervision and Curriculum Development, NEA, 1969, 2-13.

Ulrich, R. E., Louisell, S. E., & Wolfe, M. The learning village: A behavioral approach to early education. Educational Technology, in press.

Unikel, I. P., Strain, G. S., & Adams, H. E. Learning of lower socio-economic status children as a function of social and tangible reward. Developmental Psychology, 1969, 1, 553-555.

Walker, H. M. & Buckley, N. K. The use of positive reinforcement in conditioning attending behavior. Journal of Applied Behavior Analysis, 1968, 1, 245-250.

Wasik, B. H., Senn, K., Welch, R. H., & Cooper, B. R. Behavior modification with culturally deprived school children: Two case studies. Journal of Applied Behavior Analysis, 1969, 2, 181-194.

Webb, J. N. Taxonomy of cognitive behavior: A system for the analysis of intellectual processes. Journal of Research and Development in Education, 1970, 4, 23-33.

Weber, G. & Marmion, W. H. Merit pay and alternatives: Descriptions of some current programs. Washington, D.C. : Council for Basic Education, 1969.

Weissman, R. Staff differentiation. CTA Journal, January,1969, 40-45.

Winter, S. K., Griffith, J. C., & Kolb, D. A. Capacity for self-direction. Journal of Consulting and Clinical Psychology, 1968, 32, 35-41.

Wynne, E. Motivation for change: Output measures. In J. Guthrie and E. Wynne (Eds.), New models for public education. Englewood Cliffs: Prentice Hall, 1971.

Zangwill, W. I. The achievement grant program. Unpublished manuscript, Department of Health, Education, and Welfare, Washington, D. C., 1970.

APPENDICES

APPENDIX A

DOCUMENTATION OF EDUCATIONAL PROGRAMS UTILIZING INCENTIVES TO STUDENTS

Behavior Analysis Program for Follow Through

Director: Don Bushell, Jr.

Site: Department of Human Development, University of Kansas, Lawrence, Kansas.

Subject Population: Culturally deprived children enrolled in Follow Through classrooms, grades K-3.

Description:

The Behavior Analysis program has been designed to meet the needs of the culturally deprived child. In order to do this the program must identify and supply the basic skills which are needed in the classroom and design reinforcement procedures which are effective where those normally existing are not adequate. The Behavior Analysis program has taken the basic principles of human learning out of the laboratory and put them to work in the schools.

The first step in implementing the system is to define an instructional objective. The program emphasizes development of social behaviors as well as the basic academic skills of language, reading, writing and mathematics. The second step is to determine how much the child already knows about what is being taught. Each child differs in the skills that he brings into the classroom so that it is necessary to decide where each individual should begin working. The final step is to establish an effective system for classroom reinforcement so that the child is well motivated to learn.

Token Economy. A token economy system is used so that behavior can be reinforced immediately without disrupting the on-going classroom activity. Also, the child can then decide to exchange tokens for activities which he values. Potential incentives in every classroom include recess and games, stories, snacks and the chance to select a favorite activity. These events, when made available as a direct consequence of desired behavior, act as powerful reinforcers to generate and sustain motivation and progress. As skill increases, fewer tokens are necessary to support progress.

In order to provide frequent attention and individual reinforcement, parents are used as aides in the classroom. The professional teacher heads the team and is assisted by a full-time aide. Two parent aides concentrate

on individual tutoring. Training of parents consists of a six week program in the classroom, then that parent trains another. It is felt that parents who work in the classroom extend the benefits of the program into the home, thus becoming partners in the education of their children.

Curriculum. Almost all instruction is conducted in small groups or individually and the instructional needs of each child are individually assessed to allow him to advance at his maximum range. Instructional materials are selected from those currently available and which meet the following criteria:

(1) specify what the child will be able to do at the end of the sequence
(2) require frequent responding by the child
(3) contain clear criteria for a correct response
(4) allow for individual rates of progress, and
(5) provide for periodic testing of achievement gains.

Following an intial emphasis on developing social and classroom skills, the core subjects of reading, mathematics and handwriting are stressed.

The daily schedule of the classroom includes staff planning, formal instruction and reinforcement activities. The specific lesson plan is determined by the progress of the child but the three core subjects are taught each day. At the beginning of the year there is frequent alteration between study periods and special events but at the end of the year the pattern has reversed so that 10-15 minutes of special activity supports 45-50 minutes of work.

The general strategy for classroom discipline is to ignore inappropriate behavior and provide frequent reinforcement for desirable behavior. Neither verbal nor physical coercion are ever used. In cases where the behavior is damaging or dangerous, a time-out procedure is used. The child is told what rule he has broken and is seated in a chair away from the rest of the class. A timer is set for three minutes at which time he returns to the group. Punishment consists of the fact that there has been no opportunity to engage in behavior that results in token reinforcement.

Program Implementation. The development of a Behavior Analysis program occurs in three phases. University of Kansas is initially responsible for the design in all characteristics and details. Phase one comprises the first year or two of the program in which the sponsor provides regional, district, and individual classroom training with it's own staff. Workshops and an in-service

course in principles of behavior modification are among the procedures used to support program implementation. During this phase two administrative positions are developed to manage the program at the local level. The program coordinator has duties similar to a principal in being responsible for coordinating all elements of the program. The parent coordinator recruits parent aides, provides orientation and training for parents and fosters self-help activities among Follow Through parents.

During the second phase, local leadership assumes more of the training and support responsibility. Staff Training Coordinators, usually Follow Through teachers, are expected to be expert in the methods and procedures of the Behavior Analysis program and skilled in teaching these to other teachers, aides and parents.

The third phase begins as the first group of children complete the third grade. By this time, local training staff, experienced teachers, aides, and parents are able to continue the program with only periodic consulting by the University of Kansas.

Training. To accomplish the objectives defined by the Behavior Analysis program a series of workshops are scheduled throughout the year. Twice a year, a professional level seminar is conducted at the University of Kansas and attended by the staff of the Kansas Support and Development Center, the training coordinator from each district, and training and curriculum consultants. All current program procedures are discussed and common problem-solving strategies are developed.

One week regional workshops are scheduled in August attended by a delegation of teachers, aides, and training coordinators from each district. These delegates subsequently serve as leaders in the semi-annual workshops. Two three-day workshops are held in each district. Emphasis is given to teaching and reinforcement formats, changes in curriculum materials,and roles of classroom personnel. One day workshops are conducted in each district in October, November, February, March and April. These sessions insure that problems never last more than several weeks before being corrected.

Evaluation. There are five basic procedures used to evaluate the effectiveness with which each classroom is implementing the Behavior Analysis program. At the end of each week forms from every classroom describing each day's activities are mailed to Kansas. By noting the beginning and ending time of each event together with a description of the procedures followed, information is provided which indicates how explicitly Behavior Analysis procedures are

being followed. These records are also used to provide feedback to the teacher. Also sent at the same time are individual progress records for each child noting status in reading, writing, and mathematics. These rates of progress indicate the need for differential procedural changes. It is felt that these data provide more detailed information on pupil achievement than standardized achievement tests. Video tapes are made on a regular schedule portraying lessons or behavior which a teacher would like to modify. These are analyzed by a team at Kansas and recommendations are made to individual teachers about needed program improvements.

Pre and post year standardized achievement tests are given to selected classes and control groups. This is a means of clearly documenting progress of Follow Through children. In addition to evaluations conducted by local districts, several Behavior Analysis districts are included in a sample which Stanford Research Insititute has drawn for purposes of national evaluation.

Behaviorally Engineered Environments

UMREL - Upper Midwest Regional Educational Laboratory

Director: David N. Evans

Deputy Directors: Marvin F. Daley, William M. Ammentorp, William J. Hill

Site and Subjects: The program is being developed in twenty demonstration contingency-managed classrooms at four sites. Two of these, John Hay Elementary School (4 classrooms) and St. Stephen's Elementary School (8 classrooms) are in the Minneapolis inner city. Seven classrooms are located in Minnesota's Red Lake Reservation Elementary School and there is a kindergarten in Elk Point, South Dakota.

Description:

The primary objective of UMREL's program is to design and implement a behaviorally engineered educational environment where the primary focus is the individual student and the basic dependent variable is his academic performance. The behavioral management program is based on scientific principles of stimulus and response manipulation and reinforcement of desired performance. An integral part of the program development is controlled observations of events and their surrounding environments. This data is used to design procedures which have a demonstrable effect on the student's behavioral repertoire. Components of the educational process, including curricula, classroom ecology, teacher training,and evaluation, have been subjected to systematic analysis and redesign. Innovations in these areas are discussed below.

Behavioral Objectives. Individualized instruction is the keystone of the curriculum. Instructional objectives are divided into microtasks programmed in sequential order and specified in behavioral terms. Behaviorally written instructional objectives provide a measurement index for pupil progress evaluation and serves as a guide for teacher decision-making.

To facilitate lesson planning,an Individualized Lesson Format guide has been developed. A major concept is identified and then restated in operational descriptions of the educational objective. These include action - statement of the learner's response; measure - the frequency, duration or quality of the response; condition - statement of the conditions

under which the response will occur; and, criteria - establishment of the performance level at which the student must respond. The teacher then chooses the materials needed based on the specified activities. The final section of the guide provides for teacher evaluation of student progress. The learner's response is the measurement index used when evaluating methodology and materials. If students fail to meet the objectives, lessons can be reformulated based on the data collected.

Behaviorally Engineered Classroom. The Premack principle, which states that any high probability behavior can be used to reinforce a low probability behavior, is utilized in arranging the classroom environment. The room is organized into a task area and a reinforcing event (RE) area separated by a progress check point. The teacher assigns each child a lesson and circulates around the room helping children when needed. When finished with his work the child proceeds to the progress check point to have his work checked by an aide. If he has made errors, he is complimented on the work that is correct and sent back to his desk to correct the errors. If his work is 100% correct the child is allowed to choose an activity and play in the reinforcing event area for five minutes. At the end of this time the aide asks him to return to the task area and he is given another assignment.

Activities from which the child may choose are pictured on a large poster board called the reinforcing event menu. In order to determine reinforcing events a questionnaire was developed and administered to the classes regarding favorite games and interests, use of free time in school, and spending of money. Such activities as games, art and crafts work, radio, records, animals and food were identified. It is felt that the effectiveness of the reinforcing event area is a function of the type and number of events available within classroom restraints.

Teacher Training. Teachers in the developmental classes were given both pre-service and are currently engaged in in-service training primarily by means of laboratory workshops and seminars. Training focuses on skills of contingency management, principles of behavior objectives, and behavioral analysis of learning problems.

A programmed instruction training project was developed which included: (1) a programmed text, (2) videotaped samples of student behavior for practice in applying contingency management principles, and (3) a series of progress checks with which the teacher can evaluate her performance in the program.

The ultimate criterion for the training program was change in the student teacher's behavior. Specifically, she was required to:

a) Select and write instructional objectives for each curriculum area.

b) Measure and record specific student responses.

c) Transfer data from a recording sheet to a graph and interpret behavioral changes.

d) Specify present level of task performance and increase on a systematic basis.

e) Establish performance criteria for task behavior.

f) Select high probability behaviors of a student based on observation and measurement.

g) Design a classroom invironment to include Task Area, Progress Check Point, and RE Area.

h) Arrange a contract including reinforcing events which will be presented contingent on specified task completion.

Several reinforcement programs for teachers and aides have been developed. In-service educational credits, applicable to salary increases are being used as reinforcement for specific teacher behavior that affects student performance. In other words, teachers are evaluated by and reinforced for increases in student performance. Presently, the course consists of three componenets, for each of which one in-service credit may be earned. It is hoped to increase this level of credits as new objectives are added to the UMREL program.

Another incentive project for teachers and aides was designed to increase the accuracy level of filling out pupil performance data forms for machine and computer use. A total sum of $80 was allocated to each class and delivered in increments based on specified accuracy level contingencies. For example, 50% of a class' forms at 100% accuracy would earn $20. The only restriction was that the money be spent in some way to benefit the class, this decision being controlled by the teacher. Results of this project showed that the use of money as a reinforcer produced mixed reactions. Some teachers resented receiving money for work which was "part of their job", yet others were enthusiastic about it. Another problem was in the delivery system. Reinforcement oftentimes was not immediate enough to be effective.

Classroom aides are felt to be necessary in order to free the teacher to teach. The aides are central to the process of recording pupil performance data and monitoring the RE area. In order for the use of teacher-aides to be effective , they must be trained in specific objectives which result in measurable behavioral outcomes. To this end UMREL has developed a programmed

training manual for aides which concentrates on skills and knowledge required for the specific tasks they will be assigned.

Data Collection. One of the basic means of collecting data is the Pupil Performance Data Form (PPDF) which is designed for computer use. There is one form for each curricular area per child per week. Detailed information is recorded on task level, time consumption, errors, evaluation and RE choice. Up to 1200 items of data per student are recorded by teachers and aides each week. This allows for weekly feedback on the number of reinforcements a student earned (i.e., number of tasks completed), number of pages completed, number of errors prior to mastery, etc. With this information the teacher can evaluate an individual's program and make necessary adjustments.

Another data collection technique is the use of videotape. Videotaping is used both as a behavioral observation tool and a method to teach behavior observation. An audio track paced to the videotape,instructing the observer in which student to watch, how long and what behaviors to record,ensures reliability of observations. In addition, observations can be made at any time by any number of observers.

Also developed and being expanded is a storage and retreival bank for sequentially arranged instructional objectives. The objective gives the teacher a precise statement of what is expected of the student and provides instructions for materials and procedures for evaluating student performance. A second division of the bank contains data concerning the effects of various procedures on behavior. Practicing educators are able to query the bank about the most effective and efficient methods to reach specific educational objectives.

CASE - Contingencies Applicable to Special Education
Institute for Behavioral Research, Silver Springs, Maryland

Director: Harold L. Cohen

Site: National Training School for Boys, Washington, D.C.

Subject Population: Sixteen (CASE I) juveniles adjudicated in a penal institution. The ages ranged from 14 to 18 and comprised a racial and state regional balance. All were school failures, 85% being school dropouts with the average grade level completed being 7.84. The average length of sentence was 2.52 years and all but 3 had a history of institutionalization.

Description:

CASE I was a six month pilot developmental study toward the two year CASE II - MODEL (Motivationally Oriented Designs for an Ecology of Learning), twenty-four hour control project. Although the research model developed was specific to a penal institution, a further aim was to produce a curriculum and educational technology generalizable to other educationally and culturally deficient adolescent populations. Most of the material in this report is taken from CASE I data but the procedures used, although expanded, were essentially the same for CASE II.

The objective of the CASE project was to develop methods to maintain educational behaviors in the student-inmates. The most general purposes were defined by behavioral patterns most notably lacking in the boys which included (1) shaping attitudinal behaviors to more nearly approximate those established in the public school system, and (2) to raise a specified list of subject matter performances to a level approximating their age/class group in the public school system. Continual measurement of educational behaviors was built into the project to serve two major purposes. As learning was continually measured and recorded, the staff was able to evaluate the efficacy of the procedures used. Also, measurement acted as feedback to the boys and allowed for learning to be translated into specific behaviors.

Behavioral Procedures. A system of extrinsic reinforcement utilizing points given for desirable educational behavior and exchangeable for valued privileges and activities was used. Points were earned only by academic performance and were accumulated in a banking system. Earnings were used to purchase soft drinks and snacks, time in a recreation lounge, office space, books and magazines, and mail-order material (each point worth $.01).

It was found that social praise and being correct in academic achievement were strong reinforcers for maintaining performance. The boys continued working at their lessons even after they had accumulated enough points to buy time in the lounge. It was also found that the status of working on higher level courses was a powerful reinforcement to complete the prerequisites.

The program was voluntary and individualized. The student-inmate could convert his points into any specified reinforcer or save them if he chose. He worked at his own pace on individualized curricula. A grade of 90% correct on programmed work was required to earn points.

Engineered Environment. To establish the conditions of learning, measurement, reinforcement and point-conversion, a special environment was designed based on a behavioral approach to the use of architecture and equipment. Private student offices were constructed to provide an area for study behaviors and teaching machines. They were also used as reinforcers - required to be "rented" by maintaining a specified level of academic performance.

A lounge was established with a juke box, TV, vending machines and other activities which appeal to teenagers. Eventually extensions were made which included a quiet game room (i.e., chess, word puzzles, etc.) and a library for reading newspapers, magazines and books. Everyone was allowed into these areas as long as he fulfilled the scheduled prerequisites and paid the appropriate amount of points. In addition, there was a store which, by recording types and quantity of items purchased, became the principal means of according value to the conditioned reinforcer - the point.

Programmed Instruction. CASE's basic premise was to incorporate the principles of environmental planning and behavioral psychology into a program enhancing students' academic behaviors. By Stanford Achievement Test measurement, the students had a range of academic ability from the second to seventh grade and materials were provided consistent with these abilities. The basic learning materials were linear programmed instruction courses allowing for differences in rates of learning. These courses were employed both as total subject matter units and as prerequisites for live teacher classes (i.e., lectures, seminars,

workshops and laboratories). Although subject matter covered included a variety of elementary and junior high school subjects, no attempt was made to structure a curriculum. The goal of the project was to find a means of rehabilitating juveniles in educational procedures. Standard classroom activities were also planned to supplement programmed instruction but each classroom activity had some educational prerequisite for registration. All students who wished to take these courses were required to pay a registration point fee.

The type of record keeping found necessary included time spent in programs and classes; frames of programmed material completed; rate of correct responses; points assigned for each activity; special instructions; and pre- and post-test results. With these data recording functions it was possible to constantly monitor the students' progress through all educational work.

Evaluation. The basic requirement for the students was an ability to use programmed instructional materials, loosely defined as a fifth-grade reading ability. Using reading skills as a criterion of measurement and relying heavily on the Gates Reading Survey, the group average was found to be 5.6 grades with a range of 2.5 to 7.3. Results of the Stanford Achievement Tests administered at the time of the students' entrance to the National Training Institute corroborated these achievement levels. These entrance testing scores were employed for evaluation of grade level change in the pre- and post-testing batteries. The entire battery was composed of twenty-one instruments including traditional achievement tests, programmed instructional material tests, and others specifically developed for entrance assessment of student behaviors.

The average subtest grade level results of the Stanford Achievement Test in the pre-post battery were as follows:

(a) Language 4.4 (pre) to 5.0 (post)
(b) Spelling 4.7 (pre) to 6.0 (post)
(c) Arithmetic 5.3 (pre) to 6.5 (post)
(d) Reading 5.5 (pre) to 6.6 (post)

The gains apparent in these scores were accomplished with three one-hour voluntary study periods available five mornings a week for six months.

Data were also collected on the amount of time students spent in various activities, i.e., educational, leisure or "free" facilities (bench outside classroom or bathroom). Given these choices, a subject's decision would indicate the strength of the several reinforcers. Manipulation of the basic

contingencies was clearly reflected in work and lounge activities. When point values required for entrance ot the lounge were raised, amount of time spent in educational activities increased. The opening of tutored classes also provided an increase in time spent on educational activities.

The increase in educational behaviors and change in attitudinal behaviors encouraged the CASE staff to expand the project into a twenty-four hour learning laboratory. It was also evident that the procedures used were not only effective in generating and maintaining student growth but that they could be taught and implemented in an existing institutional setting. Essentially, the CASE mode of operation relies on immediate reinforcement and delayed punishment rather than the reverse which is the standard penal system.

Juniper Gardens Children's Project

Director: R. Vance Hall

Site: Juniper Gardens district, Kansas City, Kansas. The project is conducted under the auspices of the Bureau of Child Development, University of Kansas, Lawrence, Kansas. The neighborhood where the experimental schools are located and from whence the children are drawn, is an outlying poverty area of Kansas City. Housing is old and delapidated but conditions do not equal that of an inner-city tenement slum.

Subject Population: Since its inception in 1965, there have been a number of children enrolled in the Juniper Gardens schools, varying from year to year and as a function of the research being conducted. Ninety per cent of the subject population are Black and are considered culturally deprived. About 50% are from homes with only one parent.

Description:

The Juniper Gardens Project originally consisted of two experimental preschools. The Turner House program is aimed at hard core poverty children and is designed to teach them skills that will be useful in public schools. The Parent-Teacher Cooperative program is for members of the "upwardly mobile" poverty class and is directed toward getting parents involved in the educational process. Mothers are taught how to modify the behavior of their children and to teach them skills they will need in school. Each of the experimental schools are staffed by advanced graduate students under the supervision of Dr. Hall.

The procedures used are a strict behavior analysis approach based on operant conditioning theory. Emphasis is placed on defining desired responses in behavioral terms, rigorous observation and measurement techniques and the use of many types of reinforcers. The general research paradigm is an ABAB design; i.e., baseline, treatment, reversal, and reinstatement. Measurement techniques include time sampling and recording the rate of response and the use of videotape. Rater reliability is often checked by using two independent

observers. Data is tabulated in graphic form to concisely convey the results of treatment.

Types of reinforcers and conditions of reinforcement delivery are experimentally manipulated to achieve the desired behavior. It was found that with extremely deprived children, warm praise was rarely powerful enough to effect extensive changes in behavior. Material reinforcers such as snacks and candy are initially used to reinforce low frequency behaviors, always accompanied by statements of approval and affection. In this way the effectiveness of social reinforcement is increased. Another class of reinforcers that proved highly effective were the equipment and materials of the classroom. These items were therefore only made available contingent on desired behavior.

It should be noted that these techniques deal with small increments of behavior, immediately reinforcing instances of the desired response. This is especially important when shaping a low frequency of occurrence behavior. The emphasis is on specific, immediate reinforcement for small achievements.

A teacher's behavior, as well as a child's, is measured and recorded in relation to the response she is trying to elicit and maintain. This is particularly seen in the Parent Co-op program where mothers are being taught to manage their children's behavior. It was initially observed that the mothers were using almost no praise and were punishing errors with nagging or threats. Under conditions of recording and flashing a light when praise behavior occured, the rate increased significantly. When the child's responses were measured concurrently, desired performance was seen to vary directly as a function of amount of social reinforcement dispensed by the mother. With establishment of a high rate of praise and desired behavior, aversive control almost disappeared. The success of this program can be inferred from the marked change in IQ scores. For 30 beginning children the average score was 69 with a range of 48 to 81. At the end of a one year program, the average score was 87 and the range was 67 to 117.

In addition to these two experimental schools, the group at Juniper Gardens have conducted pilot research projects in many public schools in and around Kansas City and Lawrence. Two years ago they began working in the Juniper Gardens public elementary school. Teachers were instructed in classroom management and measurement techniques and natural reinforcers in the environment were arranged to achieve desired behavior. Another example is the use of a point system (exchangeable for privileges and activities) to increase study

behavior and decrease disruptive behavior in a junior high school. Results of their work, applying behavior analysis in normal and special education classrooms, is well documented in the literature. The conclusion drawn from this research is that such procedures are overwhelmingly effective in modifying behavior.

Teacher Training. Although much of the work conducted in the classroom has been under the supervision of an expert in reinforcement techniques, it is necessary for the teacher to learn these procedures in order to maintain improved behavior. It has been noted that although initially the teachers experience reinforcement from the researchers in terms of attention and social praise, the improvement in student behavior is enough to maintain the teacher's classroom behavior which had brought about the change. In other cases, the teacher discontinues experimental procedures when the research is concluded.

In order for a teacher to become skilled in using behavior management techniques, it is deemed highly desirable to have them enroll in workshops or college courses devoted to this purpose. To this end, Dr. Hall has developed a program which includes (1) basic learning theory principles, (2) experimental studies in the literature, and (3) techniques of defining, observing and recording behavior. Other objectives are to have the student carry out a study using experimental procedures and to give him contact with studies carried on by other class members. The course meets several times a week over an entire semester and the students earn academic and professional credit. The success of this program is indicated by the growth in enrollment which has increased from nine to a maximum of 50 per class in two years. Additional sections are being added to meet the increasing demand from teachers, principals and counselors.

Learning Intervention Systems

Director: Dan Meyerson

Site: Slater Elementary School, Mt. View, California

Subject Population: There are 480 students at Slater Elementary School, comprising grades K-6. The racial make-up is as follows: 14% Mexican-American, 5% Black, 5% Oriental, 5% other East Asian, and 71% White. The socio-economic level of the families is characterized by 15% on welfare or farm laborers, 35% semi-skilled or skilled laborers, 25% lower to middle-middle class and 25% upper middle class, due to the large concentration of military families in the area.

The achievement level of the students fit a bimodal distribution with a large cluster performing approximately a year below grade level and a large cluster a year above grade level.

Description:

Various learning intervention systems based on behavior modification theory are employed in this elementary school. The models are tailored to meet the needs of individual classes and to suit the instructional orientation of each teacher. Systems are flexible and dynamic in that they respond to changes in academic and behavior problems and the contingencies are applied to either groups or individuals or a combination of both. Intervention procedures include the following:

(1) Individualized Social Reinforcement System. (Kindergarten)
While lead teacher conducts large group activity, assisting teacher looks for and gives personal attention to small increments of improved performance.

(2) Self-Programmed Teacher Reinforcement.
Teacher is self-programmed to continuously recognize and socially reinforce increased competencies of individual students. Each class member's appropriate behavior is reinforced several times each period through teacher attention.

(3) Contract Reinforcement System.
Teacher has a series of contracts with pupils who have exhibited learning and behavior problems. Reinforcement of appropriate behavior with "points"

by teacher or student leads to contracted privilege. Inappropriate behavior is extinguished by ignoring or time-out procedures.

(4) Engineered Classroom.

Pupils with severe learning or attention problems are reinforced with tokens leading to privileges as they work at the achievement center. If they are unable to work at the achievement center, they continue to be reinforced at one of the seven other learning centers around the room.

(5) Student-Operated Token Economy System.

Students work for "chips and checks" which they use as currency to purchase privileges which they have chosen and to which they have ascribed "chip" values. Teacher reinforces recitations and completed work with bonuses for creativity and increased quality of performance.

(6) Playground Supervision.

Noon aides award tickets randomly for helpful and appropriate play behaviors. Pupils cash in tickets for a surprise treat at variable unannounced times.

(7) Student Tutors.

Fifth and sixth grade students, trained to recognize and acknowledge adaptive responses while ignoring maladaptive behaviors, teach specific micro-lessons to younger pupils.

(8) Parent Education Program.

Parents meet periodically with their child's teacher to analyze child behavior and develop strategies to institute new behaviors based on behavior modification principles. The parents are also included in contracting arrangements made with the child.

The common elements of the various systems are the emphasis on reinforcing desired behavior, lack of aversive control (except use of time-out in extreme behavior problems) and the general criteria by which the systems are evaluated.

An example of specific techniques that are used involves two boys who were described by their teacher as being "completely out of control." It was decided to videotape the class during a multimedia art period and focus on the two hyperactive boys in order to obtain baseline behavioral data. From this point an individualized program would be started to modify the undesired behavior. As it happened the two were engaged in their art activities 100% of the time during the videotaped session. The film was shown to the classroom, noting that the two boys were engrossed in their activity. The individualized program was not required as the two boys immediately afterward responded to

instructions and attended to their work. This was explained as a variant of modeling. Evidently, seeing themselves on tape engaged in task-oriented behavior was sufficient reinforcement to effect a complete change in their behavior and self-image.

Another example involves 12 children in the fifth grade who were reading two grade levels below the norm. Dr. Meyerson works with these children in groups of four for a period of twenty minutes in the morning, dispensing rewards for specifically stated reading behaviors. If a certain level of performance is achieved, the child is allowed free time in the library to read whatever he wishes. The childrens' reading performance has improved considerably and it appears that they will have advanced one grade level in a semester. It is hoped that they will be reading at the norm by the beginning of the 6th grade.

Teacher Training. Training in techniques and procedures is continous and adapted to each teacher's needs. A workshop, attended by all teachers, is conducted quarterly to train, advise and reinforce the teachers' reinforcing behaviors. In addition, they are urged to attend conferences and seminars on behavior modification in the classroom. All of the teachers are required to take a course in the theory and applications of behavior modification offered through San Jose State College extension division. Many of the teachers worked as student teachers under Dr. Meyerson. Behavioral data indicate that, on the average, teachers are rewarding desired responses 50% of the time.

Parental Involvement. Efforts are made to reach the parents of all the children and explain the systems which are being used in the classroom and to persuade parents to institute these procedures in the home. Such activities as parent-teacher workshops providing a short mini-course in behavior modification techniques and potluck dinners and other social affairs have generally generated a low rate of attendance. This is explained by several factors: (1) over one-third of the children have only one parent at home who works (typically semi-skilled or skilled labor) and who does not have the time or energy to attend such functions, although they express interest;(2) many children come from a lower class Catholic family with an authoritarian tradition, i.e., whatever the school does is alright with them, and; (3) a large minority of children are from mobile, military families who do not tend to form strong involvements with educational and community systems.

A procedure which is seen to be most effective is to arrange coffee sessions at a parent's house, encouraging her or him to invite other parents. Dr. Meyerson and/or the teachers then meet with them in these informal group sessions in which the parents are more comfortable and communicative. The majority of parents are responsive and positive toward reinforcement techniques but there are some who maintain that children "should not be rewarded for work which they are suuposed to do."

Dr. Meyerson feels that the traditional attitude of educators of blaming academic failures on home situations which they cannot change is a fallacy. The school environment is unrelated to the home and given the proper techniques learning can and does occur.

Evaluation. The two general criteria by which the various intervention systems are evaluated are (1) the extent to which academic and behavioral objectives are met, and, (2) the extent to which teachers and students respond favorably to the learning environment. It is considered of primary importance that both of these criteria be met; not just one or the other. Behavioral objectives are stated in explicit terms, i.e., units of work accomplished or percentage of correct responses required for a reward, and performance rates are recorded in terms of these objectives. In general, reinforcement procedures have increased performance levels of most students in all subject areas. Some students have improved scores on arithmetic and spelling tests 400%. In addition, Dr. Meyerson reported that before instituting learning intervention programs he spent 60% of his time counseling troublesome behavior students. He now spends only 1% of his time attending to discipline problems.

The Learning Village

Director: Roger Ulrich

Site: Kalamazoo Valley Intermediate School District, Michigan. The Learning Village is a private school sponsored by the Behavior Development Corporation. This program grew out of the work at the Behavior Research and Development Center at Western Michigan University.

Subject Population: Divisions of the school include the infant nursery which enrolls children aged two to thirty months; the nursery program (2 1/2 to 5 years); and the grade school (5 to 11 years). The children come from a wide range of economic and cultural backgrounds, many being described as academically disadvantaged.

Description:

The purpose of the Learning Village is to accelerate and enhance the development of children by applying scientific principles. The methodology of behavior modification is thought to be especially appropriate to educational settings since one of the principal concerns is with the acquisition of new behaviors. An important group of these behaviors is academic performance including effective use of language and abstract concepts and the acquisition of information. Another group is the personal behaviors and understanding the causes of one's own behavior. The most important group is seen to be those which comprise the social and emotional responses required for the individual's survival in society.

Occupying private status, the Learning Village allows for a rigorous application of behavioral methodology to all aspects of education and to children of a wide range of ages. It receives financial support from federal, state and local agencies as well as tuition from some parents. The tuition of many infants and nursery school children is paid by the Michigan Department of Social Services.

Curriculum. Children attend school all day and on a year 'round basis, spending a substantial portion of their time learning academic material. It is felt that structured early education is important and should become more widespread. The infant program is designed to teach motor, perceptual and

conceptual skills. Emphasis is also placed on personal and social skills including responses to other children and adults. The staff-to-student ratio in the infant nursery is one to three thus providing more attention and a richer social environment than that found in the typical home. Study periods are semi-formal, lasting about 10 minutes and are individual. Although data on infant day care centers are scarce, it is felt that a well designed and well staffed group day care program is preferable to a deficient home environment or care by a single "babysitting" agent. It is hoped that as the effects of such programs become known the number of centers will increase.

The personal and social goals of the nursery program are continuations of those of the infant nursery. This program expands academic situations to four 20-minute study periods each morning. These periods include language skills, reading, arithmetic, science, social studies, and the scientific exploration of the environment. During study periods, children are divided according to their current progress into groups of five each. This allows for individualized instruction while each child remains in the social context of his own age group.

At five years of age the children progress to grade school where study periods are longer and the material is more advanced. It is felt that continuation of educational techniques into grade school is necessary to maintain the gains made in preschool. The structuring of study periods is designed to make them as fun and rewarding as play periods so that the process of acquiring skills enhances the childs development.

Behavioral Methodology. The program of the Learning Village is designed to arrange the environment in such a way that the occurrence of desired behaviors will increase and undesired behaviors will decrease. The first step in the program is specifying the behaviors ultimately desired, such as proficiency in academic skills, acquisition of knowledge, affectionate social behavior, an understanding of social systems, and a good opinion of oneself. These goals are then translated into specific behaviors which constitute approximations of the terminal goals. An effort is made to match specific required responses with the child's progress to insure that his experience includes many opportunities for successful responding.

The most important tool used to promote desired behaviors is positive reinforcement. Such "rewarding" events are identified by observing the children's behavior and determining the effect of an event on the frequency with which a behavior occurs. The procedure of contingency contracting is often used

where the teacher specifies the responses required for reinforcement. Continuous reinforcement is used only to establish behaviors which have a low probability of occurrence. It has been found that intermittent reinforcement is more effective in maintaining high rates of response and is, of course, far less time consuming.

Various types of reinforcing events are used including praise, attention, opportunity to engage in a fun activity and special privileges within the school system. Social reinforcement is used lavishly and, besides being effective for most children, makes school a happy, supportive place to be. Behavioral reinforcers include toys, trips, and opportunities to run and scream after periods of quiet. A token economy system is also used to provide immediate reinforcement without disrupting ongoing activities. At the proper time, the child exchanges his tokens for activities of his choice. Deliberate, frequent reinforcement is felt to be especially important when very young children are learning simple things. As the acquisition of information acquires "intrinsic" interest, the use of reinforcers is faded out. It is always emphasized though, that the best way to make learning pleasant and exciting is to reinforce it with a wide variety of satisfying and exciting experiences.

Staff and Training. The overall staff-to-student ratio of one to five is maintained by fully utilizing the teaching capabilities of everyone involved with the Learning Village. This includes professional psychologists, certified teachers, college and high school students, parents, cooks and nurses. The procedures of behavior modification are specific enough so that all of these people can be effective teachers.

The core of the staff is made up of college students with a background in behavioral psychology and who are required to participate in an in-service training program. A program has also been instituted which identifies and trains high school students interested in education and who function as part of the teaching staff. Due to the success of these programs it is strongly felt that the education of younger students should be accelerated and integrated with experiences allowing them to make use of their knowledge and capabilities.

Another source of teaching staff is parents who have the opportunity to attend training classes in behavior modification. The Michigan Department of Social Services supports the Learning Village program of training parents in child care techniques. Besides serving as a teaching source, continuity is then provided between conditions in the home and school as parents learn to become effective modifiers of their children's behavior.

Evaluation. As of 1970 the nursery school program has only been in operation for two years and the infant and elementary programs for little more than one year. It is felt that assessment will not be adequate until the long-term development of the children has been observed.

A standardized testing program has been implemented for those age groups for which large group norms are available. The Wide Range Achievement Test is used for children in the kindergarten program. Results for 18 children, all of whom have had at least one year of behaviorally oriented instruction, show 2 reading at the fourth-grade level and 5 reading at the third-grade level. Even the lowest scorers placed well into the first grade. Arithmetic scores are a little less spectacular but more than half placed in the 90th percentile.

Data comparing the Learning Village program with a more traditional educational experience is provided by tests given in conjunction with a Western Michigan University Campus Nursery School. After one year of half-day attendance, children from both schools were given the Metropolitan Reading Readiness Test and the Wechsler Pre-School and Primary Scale of Intelligence Test Battery. Nearly all of the Campus Nursery School children then enrolled in public school. Some of the children from the Experimental Nursery School also enrolled in public school and some of them enrolled in the Learning Village. After the kindergarten year, the reading readiness and IQ tests were again administered to those children who could be located.

The post-nursery school scores show the mean IQ of the Experimental School children 7.49 points higher than the Campus School children. Although this difference is not remarkable the scores of children who subsequently attended the Learning Village kindergarten show a mean increase of 9.4 points from post-nursery to post-kindergarten. In contrast, the Experimental Nursery School children who attended public school uniformly show no change. Finally, the mean IQ score for children involved in behavioral education for two years was 10.1 points higher than that of children involved in traditional education for two years. The failure of the Experimental school children who transfered to public school to match the increase in IQ scores of the Learning Village children is felt to reflect the importance of continuing special education programs for more than one year.

On the reading readiness post nursery test, the scores of the Experimental School children are widely distributed whereas those of the Campus School children tend to cluster near the low end of the distribution. After a year of kinder-

garten, the mean percentile rank of children who attended the Experimental Nursery and public kindergarten was 80.1; those attending Experimental Nursery and then Learning Village was 82.2; and those attending the Campus Nursery and public kindergarten was 75.9. It is concluded that both the public school kindergarten and the Learning Village gave the children the needed skills to score well on the reading readiness test.

The reading performance of every one of the children in the Learning Village elementary school by far exceeds normal achievement in public elementary school. If children who attended the Experimental Nursery and are in Learning Village kindergarten continue to make progress typical of those in Learning Village Elementary, it is expected that future reading ability contrast with public elementary school children will be even greater than has been shown.

The experience at Learning Village has caused the children to learn far more than they would have from the traditional elementary school kindergarten even combined with nursery school. The staff of the Learning Village believes that these results are due in great part to the behavioral methodology used. It is felt that once the results of scientifically based educational methodology become known, growth and refinement of the technology of behavioral education should follow.

Englemann-Becker Follow Through Model

Directors: Siegfried Englemann and Wesley Becker

Site: Department of Child Development and Special Education, University of Oregon, Eugene, Oregon

Subject Population: Culturally deprived children enrolled in Follow Through classrooms, grades K-3.

Description:

The Englemann-Becker Follow Through Model is based on two major components: the DISTAR instructional materials and classroom management procedures. A combination of a specially developed instructional program and research techniques in controlling classroom behavior is felt to be necessary for improving academic performance.

DISTAR Materials. The language and reading programs have been developed from specific behavioral objectives and are particularly effective for educationally disadvantaged children and those needing remedial instruction. DISTAR Language is designed to teach basic concepts, focusing on the language of classroom instruction. Each of the 180 lessons is presented orally and involves student response, teacher's praise, correction, and repetition. Children progress or are recyled depending on their performance on frequent evaluative tests. DISTAR Reading is a two year program designed to teach children to "crack the reading code." They advance through drills in basic skills to more sophisticated activities and materials. Reading comprehension im emphasized in the Take Home Materials, which are designed to be used as rewards and incentives. DISTAR is a programmed version of the most successful techniques that came out of five years of classroom research with normal and educationally disadvantage children.

Classroom Management. Procedures used in the classroom to control behavior and enhance academic performance are based on learning theory principles and extensive research done by Wesley Becker and his colleagues. This research is well documented in the literature and includes numerous examples of modification of disruptive classrooms.

The emphasis on classroom management techniques, defining what teachers can do to make it possible for children to behave better, permits learning to occur. The first step is a teacher's recognition that how children behave is largely determined by her own behavior. Frequent use of negative reinforcement and attention to disruptive behavior is shown to maintain that behavior. Systematic application of positive reinforcement is shown to decrease disruptive behavior and increase on-task or attending behavior.

There are three basic rules taught to the teacher. The first is to reinforce desired behavior and withdraw reinforcement for undesired behavior. Heavy use of social reinforcement is stressed but will not work with all children. An important point is to reward behaviors which are both beneficial to the child's development (i.e., social and cognitive skills) and incompatible with disruptive behaviors. The second change procedure involves finding more effective reinforcers for those children who do not respond to social reinforcement. This includes token systems utilizing easily administered tokens as conditioned reinforcers exchangable for many back-up reinforcers. The final change procedure used is a combination of punishment for inappropriate behavior and reinforcement for appropriate behaviors. Punishment which involves presenting scolding or spanking is not suggested, because of such undesired side effects as avoidance. Instead, teachers are instructed to use procedures such as time-out, which remove the child from the presence of reinforcing stimuli.

Teacher Management. The job specification for these Follow Through teachers is clearly prescribed in terms of what is expected as classroom output. For instance, it is stated they will complete instruction of a certain number of lessons in a specified time period and that a certain percentage of children will successfully complete these lessons. Teachers participating in the program are made aware of the prescriptions and must agree to them. The programmed instructional materials which are used facilitate the task of job specification.

Teachers attend seminars in which they are taught the classroom management techniques and use of the instructional programs described above. A trained on-site consultant is available for support and advice. The consultant also serves an observing function to determine whether the teacher is actualy employing the techniques she has been taught. Monitoring teacher interaction is also felt to be extremely important, since teachers act as reinforcers for each other. Incentive conditions are arranged so that the teacher "doing the right thing" is rewarded and acts as a model for the others.

Family and Home Behavioral Intervention Systems

Director: Gerald R. Patterson

Site: Oregon Research Institute, Eugene, Oregon

Subject Population: Treatment is conducted with individual children who are referred to the Institute by parents, schools, courts, or other professionals. The children are classified as hyperactive, hyperaggressive, or emotionally disturbed. In all cases the child represents a serious behavior problem in the home and/or school.

Description:

The program and procedures researched and used by this group of clinicians are based on a behavioristic orientation to therapy, assuming that behavior is controlled by its consequences in the environment.

Parents, teachers, and peers are seen as important agents in the process of behavior change and maintenance. In many cases it has been found that these people are dispensing social rewards for deviant behaviors. It is felt that some kinds of deviant behaviors are acquired in social settings in which the child lacks a repertoire of behaviors necessary to secure consistent social reinforcers. In addition, even in the most non-responsive home or classroom there are behaviors (undesirable) which will elicit reliable negative reactions from the social environment. Casual observations of teacher-student interaction in classrooms suggest that few teachers provide even modest schedules of positive social reinforcement. Most of the reinforcers are highly formalized, such as grades and test scores. The control of behavior is achieved more as a function of aversive consequences. In such a system, the child who lacks requisite social and academic skills will be unable to obtain his share of the available rewards.

It is therefore seen as necessary to train parents, teachers, and peers to reduce the rate of social consequences contingent on deviant behavior and increase rates for desired behaviors. Intervention procedures for the child consist of training in competence, i.e., teaching those academic and social behaviors required to produce reliable schedules of positve social reward.

Intervention in the Home. Treatment is felt to be most effective when conducted right in the home rather than in "interview" situations in the clinic. Procedures are developed in which parents, under close supervision, produce responses that can be shaped and supported. The emphasis is on observable behavior and collection of data which determine decisions about the success of treatment outcomes.

The first step is training parents to observe. It is suggested that one reason for many difficulties is faulty tracking of the child's responses by the parents and the effect of their reactions to his behavior. The parents are first asked to specify the observable behaviors which are undesirable. They are then taught to observe and record frequencies of these behaviors. The next step is reinforcement training in which parents are taught to reduce their responses to undesirable behaviors and positively reinforce approximations to the desired behaviors.

Professional observers go into the home and record behaviors of the deviant child and interacting family members. Baseline observations are recorded for a period of six to ten days. Intervention procedures focus on the child's deviant behaviors which occur at a high rate and also on the family's responses which are maintaining those behaviors. Observations during intervention are used to provide feedback to the parents on how well they are following prescribed procedures. Techniques are revised as indicated by behavioral data.

Treatment is ususally terminated in ten to twelve weeks, but probes are made periodically over at least a six-month period.

Intervention in the School. Procedures in the school are similar to those in the home, in that some means is provided for reducing the frequency of social consequences for deviant behavior and increasing support for adaptive behaviors.

It is felt that teacher training is a necessary component of the intervention program. Although seminars or workshops may be useful, it seems important to actually get the teacher to practice the behavior in the classroom rather than just read or talk about general principles. The teacher's behavior is usually reinforced and maintained by the experimenter, particularly when he is in the classroom, but there is a problem of teacher reinforcement after the experimental program is terminated, especially if the response

cost is high. It is suggested that changes in student behavior may not be reinforcing enough and that salaries should partially be contingent on the teacher's ability to alter behavior.

The first step in a classroom intervention program is to use a "work box" placed on the child's desk. This box dispenses signals at variable intervals while desired behaviors are being observed. These signals are exchangable for small physical rewards. It has been found that a series of ten or more conditioning sessions is required to produce change which generalizes to occasions when the "box" is not present.

A particularly effective procedure which provides "training" for the peer group is to make group rewards contingent on the target individual's behavior. The deviant child then is working for the rest of the class, and his peers are rewarded for assisting his work behavior. A modification of these techniques is to individually reward peers for approach behavior to the deviant child and reward that child for his appropriate social responses.

Evaluation. In using behavioral intervention systems an emphasis is placed on precise, controlled data collection and recording techniques. Observer reliability is constantly monitored. Baseline and treatment data are analyzed in terms of rates of specific behaviors, and intervention procedures are modified accordingly. These data are also graphed to provide a clear simple means of monitoring the intervention procedures. Another important concern is the evaluation of treatment outcome in terms of professional time spent. All of the time expended on a treatment program by persons in the group, including psychologists and observers, is recorded. The goal is to refine techniques and produce maximum change with a minimum of professional input.

APPENDIX B

DOCUMENTATION OF E-Z SORT LITERATURE RETRIEVAL SYSTEM

LITERATURE RETRIEVAL SYSTEM

In order to facilitate the collection and filing of information for the literature review, an edge-punched card system, E-Z Sort, was adopted. The emphasis in such a system is on rapid recovery of information rather than storage. The use of edge-punched cards, notched for coded information, eliminates the necessity of multiple or cross indexing. All of the reference information, abstract, and multiple subject codes, are entered on one card. Entries are recovered by a mechanical operation, thus there is no need to file the cards in any order.

We have used an 8 x 10 1/2 inch card with two rows of 205 holes. As the system is expanded the use of two rows allows for superimposed indexing whereby different subject codes can be punched into the same card field. The author, title, reference, and a detailed abstract are typed on each card. Also included are code numbers which correspond to various key topics pertinent to research on incentives in education. In its present form the literature collection uses a simple one to one coding system, thus allowing for a direct sort (i.e., only one needling operation) to obtain the desired cards.

The actual mechanics of the system consist of the following:

(1) After the abstract is typed on the card, code numbers, designating relevant information in the article, are assigned.

(2) The proper numbers are then hand punched using an E-Z Sort groover.

(3) The number corresponding to the information to be retrieved is determined from the coding outline.

(4) Sorting is accomplished by inserting the E-Z Sort needle through holes in the card. Those that are desired drop out of the deck.

For example, if one wished to recover all abstracts dealing with the use of positive social reinforcement to culturally deprived children in an educational setting, three sorting operations would be performed, one for each topic.

As the collection is expanded, many more topics can be added to the coding system by using double row and multiple field indexing techniques. The present coding system outline is attached. Along with the card collection is a bibliography of every reference ordered alphabetically by author.

Therefore, particular articles, sought on the basis of author and title, can be easily located as this information is coded and punched on the cards.

Articles in the literature system came from a wide range of sources. The journals consulted (for the last five years) were as follows:

American Educational Research Journal
Annual Review of Psychology
Behaviour Research and Therapy
Child Development
Exceptional Children
Journal of Applied Behavior Analysis
Journal of Behavior Therapy and Experimental Psychiatry
Journal of Educational Psychology
Journal of Educational Research
Journal of Experimental Analysis of Behavior
Journal of Experimental Child Psychology
Journal of Experimental Education
Psychology Abstracts
Psychology in the Schools
Psychology Today
Review of Educational Research

In addition, a comprehensive search was made through the ERIC literature and relevant reports were identified. In order to secure up-to-date information on educational research projects recently completed or in progress, the staff contacted Science Information Exchange, from which it received abstracts. Complete reports were obtained from those which were felt to be most pertinent. A final major source of literature were papers presented at various conferences. Relevant papers were requested from the 1970 American Psychological Association meeting, American Educational Research Association meeting, and the regional Psychological Association meetings.

Staff plans include continuing to add to and expand the literature collection and the topic outline. This will result in a relatively comprehensive library of literature on educational research using incentives. It is felt that, due to its comprehensiveness and ease of information retrieval, this collection could be of great help of other researchers in this field.

LITERATURE REVIEW CODING SYSTEM

I Type of article
- 1 Experimental
- 2 Theoretical

II Design
- 3 Single organism
- 4 Experimental & control groups

III Setting
- 5 School
- 6 Laboratory
- 7 Home
- 8 Institution

IV Subjects
- 9 Normal middle class
- 10 Culturally disadvantaged
- 11 Disruptive behavior problems
- 12 Other special education

V Grade levels
- 13 Grades 1-3 (incl. pre-school)
- 14 Grades 4-6
- 15 Grades 7-9
- 16 Grades 10-12 (incl. college)

VI Target behaviors
- 17 Mathematics
- 18 Reading
- 19 Other academic behaviors
- 20 Social behavior

VII Reinforcing agent
- 21 Teacher
- 22 Experimenter
- 23 Peer
- 24 Other

VIII Target populations

Individual incentive placement
- 25 Teacher
- 26 Student
- 27 Parents
- 28 Peers

Group incentive placement
- 29 Teacher
- 30 Student
- 31 Parents
- 32 Peers

Problem Identifiers

Types of incentives
- 33 Material
- 34 Social
- 35 Knowledge of results
- 36 Secondary reinforcement
- 37 Vicarious reinforcement (& modeling)
- 38 Aversive
- 39 High vs low probability responses
- 40 Self-management of responses

Modes of incentive delivery
- 41 Direct, personal
- 42 Mechanical devices
- 43 Non-delivery of positive incentives
- 44 Token economies
- 45 Performance contracts
- 46 Reinforcement menu
- 47 Self-delivery

Timing of reinforcement
- 48 Immediate
- 49 Delayed
- 50 Continuous
- 51 Fixed (interval or ratio)
- 52 Variable (interval or ratio)

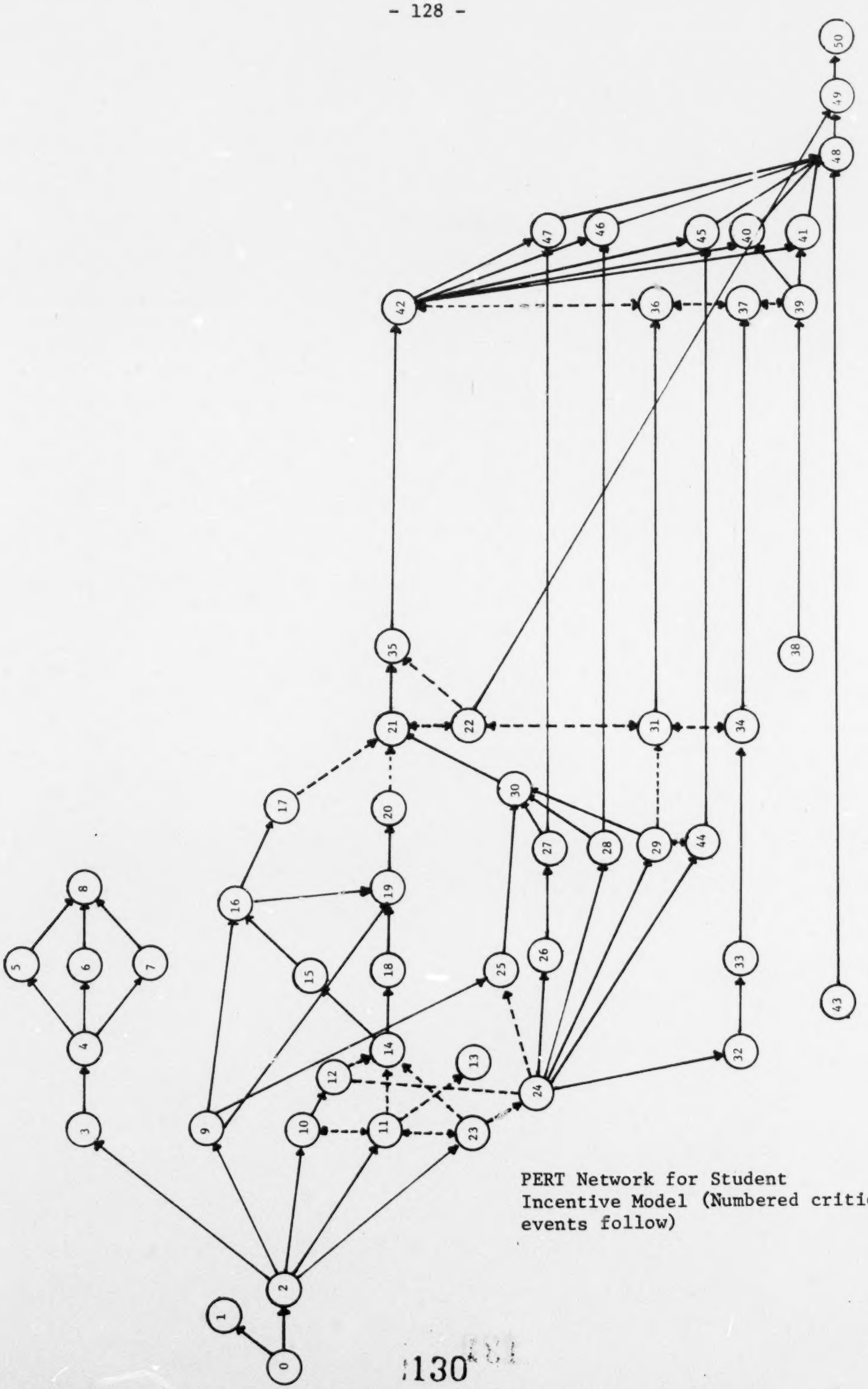
Criterion measures
- 53 Specific behaviors
- 54 Standardized tests
- 55 Criterion-referenced tests
- 56 Teacher or experimenter constructed test

Author Coding (alpha field A-Z)

first name initial - inner row
first and third last name letter - outer row

APPENDIX C

PERT NETWORKS



PERT Network for Student Incentive Model (Numbered critical events follow)

CRITICAL EVENTS IN THE STUDENT INCENTIVES MODEL

(1) signed contract received

(2) staff assignments completed
    a) on-site project staff
    b) monitoring and evaluation staff

(3) plan of public relations activities completed via discussions with school officials

(4) complete procedures and materials to explain the purposes and methods of the experiment to the general public

(5) public relations meetings with interested school board members, other interested community leaders, and potential incentive donors concluded:
    a) explanatory materials presented
    b) potential problems identified
    c) community incentive contributions obtained if possible

(6) public relations meetings with teachers' representatives concluded:
    a) explanatory materials presented
    b) potential problems identified

(7) public relations meetings with parents' representatives concluded:
    a) explanatory materials presented
    b) potential problems identified

(8) steps for diverting potential problems initiated

(9) complete selection of students who will be asked to participate

(10) range of student performance goals identified

(11) incentives for students specified

(12) performance measures identified

(13) begin arrangement for making incentives available when they are earned

(14) begin development of explanatory materials for students and their teachers

(15) explanatory materials for teachers completed
    a) overview and examples of possible incentive delivery to students contingent upon student performance
    b) range of student performance goals
    c) measures of student performance
    d) list of all measuring instruments, when each would be administered, who would administer them, amount of time required

(16) explanatory materials for teachers presented

(17) agreement from each teacher to permit classroom observations obtained

(18) explanatory materials for students completed
   a) overview and examples of possible incentives to students contingent upon student performance
   b) range of student performance goals
   c) available incentives for students

(19) explanatory materials presented by teachers to their students

(20) incentive delivery agreements negotiated by teachers with their students
   a) student performance goals specified
   b) performance measures specified
   c) formulae for calculating student incentive learnings specified
   d) incentive delivery system specified

(21) begin incentive earning period

(22) arrangements for incentives availability completed

(23) requirements for process measures identified

(24) development of performance and process measures begun

(25) begin collecting school record information on target students

(26) attitude scales construction completed for:
   a) teachers
   b) parents
   c) students

(27) attitude scales tried out, revised, and printed

(28) norm-referenced tests obtained

(29) criterion-referenced pre-tests obtained or developed, tried out, revised, and printed

(30) pre-testing completed

(31) begin periodic testing to determine interim incentive payments

(32) classroom observation scales completed

(33) classroom observers trained

(34) observer reliability checked, begin classroom observations

(35) begin earned incentive payments to students

(36) last tests for determining interim incentive earnings administered

(37) classroom observations completed

(38) begin development of self-report and interview forms

(39) self-report and interview forms developed, tried out, revised, and printed

(40) post experimental self-report forms completed by:
a) all teachers
b) students
c) parents

(41) post experimental in-depth interviews completed by:
a) all teachers
b) randomly selected students
c) randomly selected parents

(42) period of student incentive earnings ended

(43) begin development of data analysis procedures

(44) criterion-referenced post-tests selection completed

(45) criterion-referenced post-tests completed by students

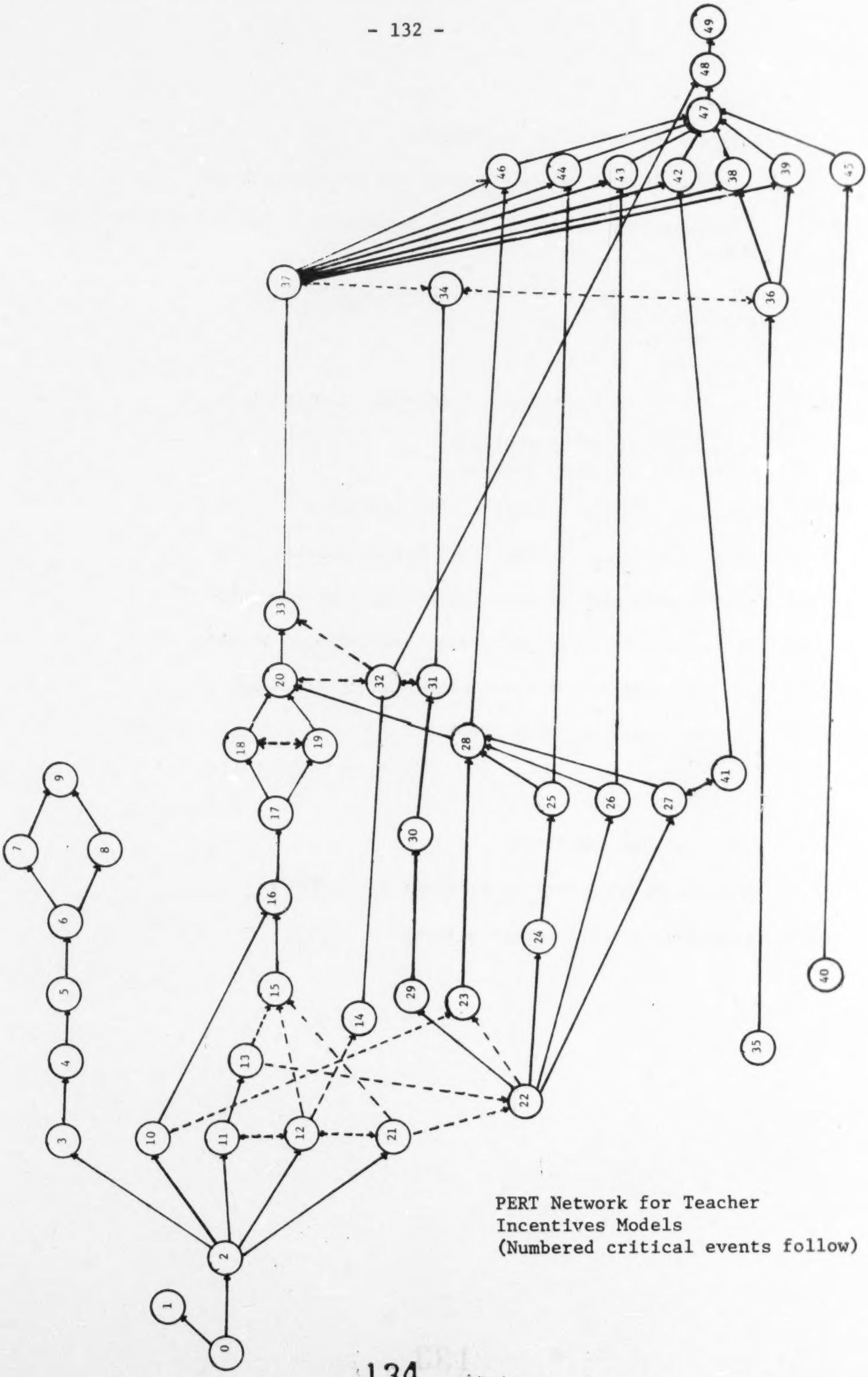
(46) norm-referenced post-tests completed by students

(47) attitude post-test completed by:
a) students
b) teachers
c) parents

(48) data analyses completed

(49) complete final incentive payments to students

(50) first year's final report written



PERT Network for Teacher
Incentives Models
(Numbered critical events follow)

CRITICAL EVENTS IN THE TEACHER INCENTIVES MODELS

(1) signed contract received

(2) staff assignments completed
a) on-site project staff
b) monitoring and evaluation staff

(3) plan of public relations activities completed via discussions with school officials

(4) complete procedures and materials to explain the purposes and methods of the experiment to the general public

(5) public relations meetings with teachers' representatives concluded:
a) explanatory materials presented
b) potential problems identified

(6) initiate steps to divert potential problems identified during public relations meetings with teachers' representatives

(7) public relations meetings with interested school board members, other interested community leaders, and potential incentive donors concluded:
a) explanatory materials presented
b) potential problems identified
c) community incentive contributions obtained if possible

(8) public relations meetings with parents' representatives concluded:
a) explanatory materials presented
b) potential problems identified

(9) initiate steps to divert potential problems identified in public relations meetings

(10) complete selection of teachers who will be asked to participate

(11) range of student performance goals identified

(12) incentives for teachers specified

(13) performance measures identified

(14) begin arrangement for making incentives available to teachers when they are earned

(15) development of explanatory materials for teachers begun

(16) explanatory materials for teachers completed
a) range of student performance goals
b) measures of student performance
c) available incentives for teachers
d) overview and examples of possible incentive delivery to teachers
e) list of all measuring instruments, when each would be administered, who would administer them, amount of time required

(17) explanatory materials for teachers presented

(18) incentives agreement with teachers negotiated
a) student performance goals specified
b) performance measures specified
c) formulae for calculating incentive earnings to teachers specified
d) incentive delivery system specified

(19) agreement from each teacher to permit classroom observations obtained

(20) begin incentive earning period for teachers

(21) requirements for process measures identified

(22) development of performance and process measures begun

(23) begin collecting school record information on target students

(24) attitude scales construction completed for:
a) teachers
b) parents
c) students

(25) attitude scales tried out, revised, and printed

(26) norm-referenced tests obtained

(27) criterion-referenced pre-tests obtained or developed, tried out, revised, and printed

(28) pre-testing completed

(29) classroom observation scales completed

(30) classroom observers trained

(31) observer reliability checked, begin observational data collection

(32) arrangements for incentives availability completed

(33) begin advance incentive delivery to teachers

(34) classroom observations completed

(35) begin development of self-report and interview forms

(36) self-report forms and interview forms developed, tried out, revised, and printed

(37) period of teacher incentive earnings ended

(38) post experimental self-report forms completed by:
  a) teachers
  b) students
  c) parents

(39) post experimental in-depth interviews completed by:
  a) all teachers
  b) randomly selected students
  c) randomly selected parents

(40) begin development of data analysis procedures

(41) criterion-referenced post-tests selection completed

(42) criterion-referenced post-tests completed by students

(43) norm-referenced post-tests completed by students

(44) attitude post-tests completed by:
  a) students
  b) teachers
  c) parents

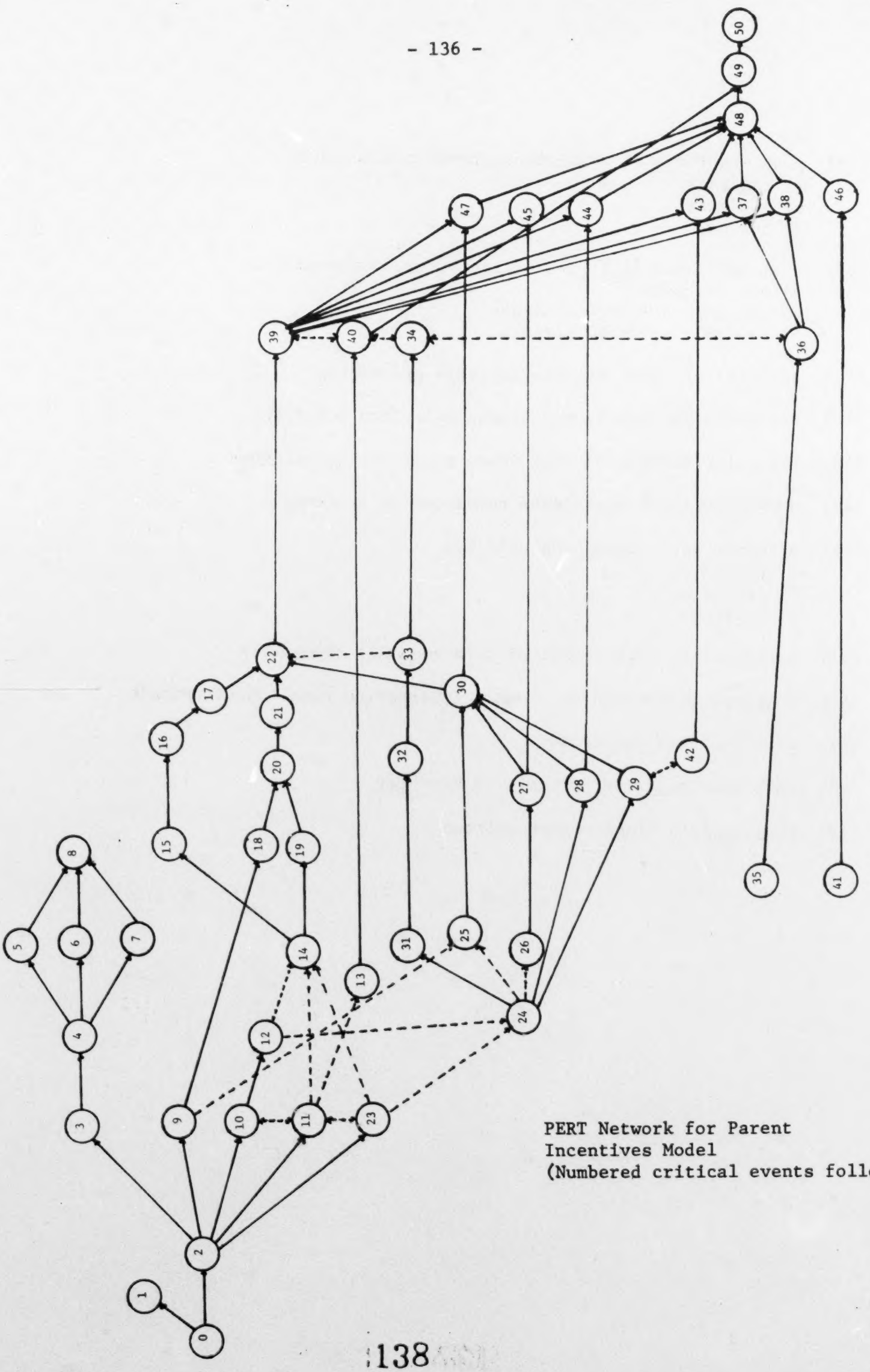
(45) complete the development of data analysis procedures

(46) complete collection of student information from school records

(47) data analyses completed

(48) complete incentive payments to teachers

(49) first year's final report written



PERT Network for Parent
Incentives Model
(Numbered critical events follow)

CRITICAL EVENTS IN THE PARENT INCENTIVES MODEL

(1) signed contract received

(2) staff assignments completed
   a) on-site project staff
   b) monitoring and evaluation staff

(3) plan of public relations activities completed via discussions with school officials

(4) complete procedures and materials to explain the purposes and methods of the experiment to the general public

(5) public relations meetings with parents' representatives concluded:
   a) explanatory materials presented
   b) potential problems identified

(6) public relations meetings with interested school board members, other interested community leaders, and potential incentive donors concluded:
   a) explanatory materials presented
   b) potential problems identified
   c) community incentive contributions obtained if possible

(7) public relations meetings with teachers' representatives concluded:
   a) explanatory materials presented
   b) potential problems identified

(8) steps for diverting potential problems initiated

(9) complete selection of parents who will be asked to participate

(10) range of student performance goals identified

(11) incentives for parents specified

(12) performance measures identified

(13) begin arrangement for making incentives available when they are earned

(14) development of explanatory materials for target groups begun

(15) explanatory materials for teachers completed
   a) range of student performance goals
   b) measures of student performance
   c) nature of parent involvement in home tutoring
   d) list of all measuring instruments, when each would be administered, who would administer them, amount of time required

(16) explanatory materials for teachers presented

(17) agreement from each teacher to permit classroom observations obtained

(18) complete arrangements for presenting information about projects to parents

(19) explanatory materials for parents completed
   a) overview and examples of possible incentives to parents contingent upon student performance
   b) range of student performance goals
   c) available incentives for parents
   d) nature of parent involvement in home tutoring

(20) explanatory materials for parents presented

(21) incentive delivery agreements with interested parents negotiated
   a) student performance goals specified
   b) some guidelines for helping students to attain those goals specified
   c) performance measures specified
   d) formulae for calculating incentive earnings to parents specified
   e) incentive delivery system specified

(22) begin parent incentive earning period

(23) requirements for process measures identified

(24) development of performance and process measures begun

(25) begin collecting school record information on target students

(26) attitude scales construction completed for:
   a) teachers
   b) parents
   c) students

(27) attitude scales tried out, revised, and printed

(28) norm-referenced tests obtained

(29) criterion-referenced pre-tests obtained or developed, tried-out, revised, and printed

(30) pre-testing completed

(31) classroom observation scales completed

(32) classroom observers trained

(33) observer reliability checked, classroom observational data collection begun

(34) classroom observations completed

(35) begin development of self-report and interview forms

(36) self-report and interview forms developed, tried-out, revised, and printed

(37) post experimental self-report forms completed by:
  a) teachers
  b) students
  c) parents

(38) post experimental in-depth interviews completed by:
  a) all teachers
  b) randomly selected students
  c) randomly selected parents

(39) period of parent incentive earnings ended

(40) arrangements for incentives availability completed

(41) begin development of data analysis procedures

(42) criterion-referenced post-tests selection completed

(43) criterion-referenced post-tests completed by students

(44) norm-referenced post-tests completed by students

(45) attitude post-test completed by:
  a) students
  b) teachers
  c) parents

(46) complete the development of data analysis procedures

(47) complete collection of student information from school records

(48) data analyses completed

(49) complete incentive payments to parents

(50) first year's final report written

APPENDIX D

LETTERS OF INTEREST FROM CONTACTED SCHOOL DISTRICTS

:143

141

dallas independent school district

Nolan Estes
General Superintendent

September 30, 1970

Dr. Steven M. Jung
Research Scientist and
Assistant to the Director
American Institutes for Research
P. O. Box 1113
Palto Alto, California 94302

Dear Dr. Jung:

Dr. Clyde Greer has referred your letter to me for response. The Dallas Independent School District will certainly be receptive to a discussion concerning future field experiments in the use of performance incentives. We are presently involved with three performance contracts with instructional firms, in addition to contracts with educational auditing firms and management support groups. We are definitely interested in increasing our present involvement.

If you would like to visit Dallas next month and discuss the project further, I will be at your service. I will be available any days except October 1-2 and October 12-16.

I look forward to seeing you.

Sincerely,

Donald R. Waldrip
Assistant Superintendent-
Accountability

DRW:gd

144

ARIZONA-MESA DIFFERENTIATED STAFFING PROJECT
966-4352 (AREA CODE 602)
DEPARTMENT OF: SECONDARY EDUCATION
ARIZONA STATE UNIVERSITY
TEMPE, ARIZONA 85281

PROJECT STAFF

FENWICK ENGLISH
DIRECTOR

JAMES K. ZAHARIS
ASSOCIATE DIRECTOR
CENTER FOR EDUCATIONAL ADVANCEMENT
809 W. MAIN
MESA, ARIZONA 85202

DAVID RICE
TOM SWEENEY
SECONDARY ASSOCIATES

LARRY FRASE
ELEMENTARY ASSOCIATE

CARMEN GAONA
SECRETARY

CONSORTIUM MEMBERS

DR. GEORGE N. SMITH
SUPERINTENDENT
MESA PUBLIC SCHOOLS
549 N. STAPLEY DR.
MESA, ARIZONA 85203

DR. JAMES OLIVERO
SWCEL

DR. NELSON HAGGERSON
ASU

DR. JOHN BELL
ASU

November 12, 1970

Steven M. Jung, Ph.D.
Research Scientist
American Institutes for Research
P.O. Box 1113
Palo Alto, California
94302

Dear Dr. Jung

This is to indicate our interest in combining our resources to investigate the potential of performance contracting in the Mesa Public Schools.

We are very pleased to have your cooperation in this endeavor. Upon his return from his worldwide tour, our superintendent will send you a letter to this effect.

We look forward to your active participation with us.

Sincerely

Fenwick English
Director

cg

FUNDED BY UNITED STATES OFFICE OF EDUCATION, EPDA P.L. 90-35 TITLE V

PORTLAND PUBLIC SCHOOLS
631 Northeast Clackamas Street / Portland, Oregon 97208
Phone (503) 234-3392
OFFICE OF INTERGOVERNMENTAL RELATIONS

Charles A. Clemans
Intergovernmental Programs
Maurice L. Caba,
Program Evaluation
Nick G. Bolos,
Program Finance
Lloyd L. Cedrus
Program Finance

October 28, 1970

NOV 0 1 1970

Dr. Stephen M. Jung
American Institute for Research
P. O. Box 1113
Palo Alto, California 94302

Dear Dr. Jung:

Jim Holmes and I found your recent visit to our district both interesting and exciting. I was most pleased to find that AIR's interest in incentives is in the very broad sense, as it is yet too early in the game to narrow into restricted areas of incentives without regard to exploring the broader picture.

You may recall that we discussed the possibility of the Portland Public Schools participating with AIR in some sort of joint venture in incentives research.

Please be assured that we are interested in such an arrangement and are anxious to further explore such possibilities with you. Such a venture might take one of a number of avenues, including the District's serving as a site for a possible research study.

Please let me know how we may further explore such possibilities with you.

Sincerely,

Charles A. Clemans
Intergovernmental Specialist

CAC:ml

SHIKELLAMY SCHOOL DISTRICT

350 ISLAND BOULEVARD
SUNBURY, PENNSYLVANIA 17801

OFFICE OF THE SUPERINTENDENT
PHONE 717-286-6763

November 19, 1970

Dr. Steven Jung
Research Scientist
American Institutes for Research
P. O. Box 1113
Palo Alto, California 94302

Dear Dr. Jung:

We are pleased to be asked to be a part of your program on Incentives in Education.

In this school system, and the others I've known, motivation to learn is a universal problem. The use of incentives to motivate students is a much needed area of exploration in the public schools. It is a need for the privileged as well as our less privileged students.

Since receiving your invitation to respond to the proposal, I've had a change in position. I have been elected Superintendent of the Avon Grove Schools in Chester County, Pennsylvania. I will assume this position as of January 1, 1971. I have discussed Avon-Grove involvement with the project with the incumbent superintendent. He thought that district would be receptive to the project; however, since he is leaving, did not want to make a commitment. After January 1 I will be able to follow up and be actively involved from Avon-Grove. We will be interested in becoming a participating school and participate in the field testing of this program.

Sincerely yours,

Harry B Gorton

Harry B. Gorton, D.Ed.
District Superintendent

HBG/em

145

Board of Education

222 MAIN STREET
WETHERSFIELD, CONNECTICUT 06109

OTTO C. HUFZIGER
SUPERINTENDENT OF SCHOOLS

OCT 29 1970

October 23, 1970

Dr. Steven M. Jung, Research Scientist
Guidance Research Program
American Institutes for Research
P. O. Box 1113
Palo Alto, California 94302

Dear Dr. Jung:

Henry Ferri referred your letter of October 8th to me and, after reading your abbreviated proposal on "Incentives In Education", I am pleased to say that the Wethersfield School Department is interested in participating in the research and development in this important area.

I am going to assign Henry Ferri, Director of Pupil Personnel Services and Special Education to this project and would appreciate your coordinating any future activities through his office.

We certainly are appreciative of your interest and we are looking forward to participating in the project.

Sincerely,

Otto C. Hufziger
Superintendent of Schools

OCH:pc
cc: Henry Ferri

APPENDIX E

SHORT INTRODUCTION TO INCENTIVES EXPERIMENT

SHORT INTRODUCTION TO INCENTIVES EXPERIMENT

The use of incentives in educational practice is universal. Incentives whether or not they are identified as such, exist for all participants in the educational process. In the present context, incentives are thought of as identifiable consequences of behavior which act to guide the future form and frequency of that behavior. Such factors as money, security, knowledge of personal success, peer or authority figure approval, fear of failure, etc., are probably operating to influence a large percentage of the behaviors which could be observed and classified in any school in the country. In this sense, the educational enterprise is not unlike other forms of human enterprise.

Recent events have stimulated serious interest in the use of incentives to improve academic performance. One such event is the flurry of contracts between school systems and private firms which bind the latter to produce specified reading and mathematics achievement gains in students in order to be paid for instructional services rendered, giving rise to the notion that outside firms know something that school personnel do not know about causing students to learn. Central to these events is the belief that the educational programs of the past decade have not produced impressive results and have especially failed the so-called "deprived" student. Whereas these failures have produced a certain pessimism in some circles, other educators have thought enough of the power of currently available techniques to venture their own capital on a guaranteed-performance-or-no-pay-basis. An examination of these techniques ususally reveals a heavy emphasis on technological innovations and "incentives" to learners.

The enclosed literature review, entitled "Manipulating Incentives to Enhance School Learning," examines research evidence regarding the effectiveness of various types of incentives in improving student performance, using various modes and schedules of incentive delivery, directed toward various identifiable incentive recipients. This evidence provides a basis for a school-based experiment which will be done beginning in May or June of 1971 to investigate conditions where positive incentives are identified and manipulated by an outside party versus the effect of conditions where incentives exist "in their natural state." The ultimate goal of this experiment is to determine the comparative benefits of positive incentives apart from the influence of major

revisions in the school curriculum. In essence, the study will investigate the effects of predetermined stipulations on the outcomes of the educational process, as opposed to changes in the materials or procedures which constitute the process. It is postulated that modifying consequences based on outcomes will lead to or require changes in methods. These changes will be of major interest; indeed, one of the major purposes of the experiment is to gather reliable "process" information regarding their nature.

The experiment is planned to extend over the course of the entire 1971-72 school year. Previous experience suggests that experiments of shorter duration cannot be counted on to produce replicable results. Moreover, it is quite likely that incentive effects are cumulative, and the differences between organized incentive and unspecified incentive conditions would become more apparent over the course of the year. Some researchers have noted that initial failure experiences produce a continuous cycle of deepening student failure. It is possible that providing identifiable incentives could cause students to gain the skills that would allow them to experience more success in school and hence become more receptive to the influence of less tangible incentives which already exist in the classroom situation. These effects should become observable in longitudinal data collected after external incentives have been phased out.

The most crucial requirement for implementing the experiment is that outputs of the educational enterprise must be made a matter of record, subject to external scrutiny. The often used term "accountability" contains the appropriate flavor. As long as teaching and learning take place in a vacuum, incentives will be extremely difficult to direct in other than a haphazard fashion. On the other hand, when the expected outputs are identified in advance and well known as such by all participating parties, incentives will follow almost as a matter of course. The AIR project staff have suggested a set of procedures which can be used to provide measurable criteria of student achievement in the areas of reading and mathematics. These are referenced in the enclosed paper entitled "Criterion Measures for Incentive Delivery."

The proposed experiment is organized according to the target populations which will receive incentives. A separate experimental model is provided for each target population. Only populations which can be

considered as legitimate entities having major responsiblity for the educational process are identified. This includes students, teachers, administrators, and parents. Outside profit-based organizations are omitted. Emphasis is being placed on disadvantaged and underachieving student populations.

In proposing a multi-site field experiment on the effects of delivery of incentives contingent upon gains in student achievement, it has previously been pointed out that important questions exist which cannot be answered without such an attempt. Basically, the questions are these: can educational goals for students be identified which are significant educationally yet provide suitable criteria for the assignment of incentives; if so, will major responsible parties in the educational process allow external incentives to be made contingent upon the attainment of such goals; if so, can incentives be identified which are effective in improving goal attainment over and above the success of existing educational programs; if so, how was this accomplished, and if not, why not?

It is interesting to note that the first two questions above must each be answerable in the affirmative before experimental methods can possibly be brought to bear to answer the latter questions. And only by embarking on an experiment can these first questions be posed.

A range of possible models has thus been constructed which could provide and evaluate the effects of incentives. School districts interested in participating in the field experiment have the option of selecting one or more of these models according to their estimates of local needs and conditions. Schools also have some options in selecting the type of incentives to be utilized, incentive payment schedules, amount of within-school competition involved, and grade level of student participants within the range 1-9. These options will be discussed further in the forthcoming final report of a feasibility study conducted by AIR which is presently nearing completion. In general, however, they are as follows:

A. Subject populations to which experimental project incentives will be delivered contingent upon student achievement gains; options include
   1. Student groups
   2. Teachers and administrators, as individuals or groups
   3. Parents as individuals or groups
   4. Any combination of the above.

B. Types of incentives to be utilized; options include
   1. Money, either for personal gain or to acquire instruction related materials
   2. Opportunity to engage in desirable activities or acquire desired materials--instruction or non-instruction related
   3. Social approval or special recognition
   4. Any combination of the above.

C. Incentive delivery schedules may be based upon various types of student achievement gains; options include
   1. Absolute group performance gains
   2. Absolute individual performance gains
   3. Group performance gains adjusted according to expected gains
   4. Individual performance gains adjusted according to expected gains.

Depending upon which options are selected, estimates for funding from external sources range form $50,000 to $150,000. The range is dependent upon the number of models a school district chooses to implement, the degree of monetary type incentives selected versus social approval of special recognition, the payment schedule chosen, and the degree to which local support may be secured. Local support would especially involve such factors as facilities, clerical and logistical backup, cost differential between the salaries of professionals involved and the salaries of substitutes employed to temporarily handle their duties while they are implementing the incentives project, etc.

Current estimates call for approximately 1800 experimental and control students to be identified at each experimental site. All incentives would be contingent upon the gains made by these students. Therefore, incentive funds requested should be sufficient to provide the maximum possible payments that could be earned; it is possible that, should maximum gains not be made, some of these funds would be returned. This is the ultimate form of accountability.